



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 144269

**TO: Suryaprabha Chunduru**

**Location: REM-2C29/2C18**

**Art Unit: 1637**

**Monday, February 14, 2005**

**Case Serial Number: 10/790430**

**From: Deirdre Arnold**

**Location: Biotech-Chem Library**

**REM 1A64**

**Phone: 571-272-2532**

**Deirdre.Arnold@uspto.gov**

### Search Notes

**RUSH**

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 363.533 Seconds  
(without alignments)  
5324.730 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183

Sequence: 1 gacgttattatgagatgagtg...aaggaagcgcgagatgacg 1183

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued Patente NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID                | Description        |
|------------|-------|-------------|--------|-------------------|--------------------|
| 1          | 1183  | 100.0       | 1183   | US-09-872-051-8   | Sequence 8, Appli  |
| 2          | 202   | 17.1        | 2378   | US-09-441-340-27  | Sequence 27, Appli |
| 3          | 168   | 14.2        | 4032   | US-09-068-101-5   | Sequence 5, Appli  |
| 4          | 168   | 14.2        | 4032   | US-09-970-921-5   | Sequence 5, Appli  |
| 5          | 147   | 12.4        | 3754   | US-09-377-4668-15 | Sequence 15, Appli |
| 6          | 147   | 12.4        | 4149   | US-09-377-4668-13 | Sequence 13, Appli |
| 7          | 147   | 12.4        | 8349   | US-09-186-002-16  | Sequence 16, Appli |
| 8          | 147   | 12.4        | 10249  | US-09-186-002-14  | Sequence 14, Appli |
| 9          | 147   | 12.4        | 10252  | US-09-186-002-15  | Sequence 15, Appli |
| 10         | 147   | 12.4        | 10339  | US-09-186-002-13  | Sequence 13, Appli |
| 11         | 140   | 11.8        | 2378   | US-09-441-340-27  | Sequence 27, Appli |
| 12         | 140   | 11.8        | 2436   | US-09-441-340-31  | Sequence 31, Appli |
| 13         | 137   | 11.6        | 1630   | US-09-441-340-23  | Sequence 23, Appli |
| 14         | 137   | 11.6        | 10846  | US-09-098-2198-5  | Sequence 5, Appli  |
| 15         | 137   | 11.6        | 10846  | US-10-164-204-5   | Sequence 5, Appli  |
| 16         | 137   | 11.6        | 10846  | US-09-923-109-5   | Sequence 5, Appli  |
| 17         | 137   | 11.6        | 10900  | US-09-098-2198-6  | Sequence 6, Appli  |
| 18         | 137   | 11.6        | 10900  | US-10-164-204-6   | Sequence 6, Appli  |
| 19         | 137   | 11.6        | 10900  | US-09-923-109-6   | Sequence 6, Appli  |
| 20         | 136   | 11.5        | 11522  | US-10-052-092-19  | Sequence 19, Appli |
| 21         | 134   | 11.3        | 563    | US-09-068-101-6   | Sequence 6, Appli  |
| 22         | 134   | 11.3        | 563    | US-09-970-921-6   | Sequence 6, Appli  |
| 23         | 134   | 11.3        | 1138   | US-09-011-151-8   | Sequence 8, Appli  |
| 24         | 134   | 11.3        | 1138   | US-09-011-151-9   | Sequence 9, Appli  |
| 25         | 134   | 11.3        | 1213   | US-09-794-384A-15 | Sequence 15, Appli |
| 26         | 134   | 11.3        | 1246   | US-09-794-384A-15 | Sequence 15, Appli |
| 27         | 134   | 11.3        | 1287   | US-08-064-121-3   | Sequence 3, Appli  |

|    |     |      |      |   |                  |                   |
|----|-----|------|------|---|------------------|-------------------|
| 28 | 134 | 11.3 | 1287 | 1 | US-08-478-015-3  | Sequence 3, Appli |
| 29 | 134 | 11.3 | 1287 | 3 | US-08-475-975-3  | Sequence 3, Appli |
| 30 | 134 | 11.3 | 1287 | 3 | US-09-084-889-3  | Sequence 3, Appli |
| 31 | 134 | 11.3 | 1287 | 4 | US-09-794-384A-8 | Sequence 8, Appli |
| 32 | 134 | 11.3 | 2345 | 3 | US-09-026-673-1  | Sequence 1, Appli |
| 33 | 134 | 11.3 | 2345 | 3 | US-09-512-650-1  | Sequence 1, Appli |
| 34 | 134 | 11.3 | 2345 | 3 | US-09-480-142-1  | Sequence 1, Appli |
| 35 | 134 | 11.3 | 2345 | 4 | US-09-573-555-2  | Sequence 2, Appli |
| 36 | 134 | 11.3 | 3544 | 2 | US-08-485-139-3  | Sequence 3, Appli |
| 37 | 134 | 11.3 | 3544 | 2 | US-08-485-139-3  | Sequence 3, Appli |
| 38 | 134 | 11.3 | 3544 | 3 | US-08-750-357-3  | Sequence 3, Appli |
| 39 | 134 | 11.3 | 3544 | 3 | US-08-750-357-3  | Sequence 3, Appli |
| 40 | 134 | 11.3 | 3658 | 3 | US-08-894-440-3  | Sequence 3, Appli |
| 41 | 134 | 11.3 | 3658 | 3 | US-08-894-440-3  | Sequence 3, Appli |
| 42 | 134 | 11.3 | 3658 | 3 | US-09-458-093-3  | Sequence 3, Appli |
| 43 | 134 | 11.3 | 3658 | 3 | US-09-458-093-3  | Sequence 3, Appli |
| 44 | 134 | 11.3 | 4032 | 3 | US-09-068-101-5  | Sequence 5, Appli |
| 45 | 134 | 11.3 | 4032 | 4 | US-09-970-921-5  | Sequence 5, Appli |

## ALIGNMENTS

RESULT 1  
US-09-872-051-8  
Sequence 8, Application US/09872051  
Patent No. 6823400  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
TITLE OF INVENTION: Coen Event PV-ZMG132(hk603) and Composition and Methods  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 1183  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(1183)  
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
OTHER INFORMATION: 165-381 construct vector DNA  
OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA  
OTHER INFORMATION: 687-1183 Zea maize genomic DNA  
US-09-872-051-8  
Query Match  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 GACGTTATTATGAGATGGGTTTATGATTAGTCCCGCAATTATACATTATACGC 60  
DB 1 GACGTTATTATGAGATGGGTTTATGATTAGTCCCGCAATTATACATTATACGC 60  
QY 61 GATGCAAAACAAATATAGCGCGCAATAGATAATATGCGCGGCGGTATCAT 120  
DB 61 GATGCAAAACAAATATAGCGCGCAATAGATAATATGCGCGGCGGTATCAT 120  
QY 121 GTTACTAGATCGGGGATATCCCGGGAATTGCGTACCAAGCTTTTAAATAGAGAAA 180  
DB 121 GTTACTAGATCGGGGATATCCCGGGAATTGCGTACCAAGCTTTTAAATAGAGAAA 180

```

Db 121 GTTACTAGATCGGGGATATCCCGGGGATTCGGTACCAAGCTTTTATATAGTAGAAAA 180
QY 181 GAGTAATTCACCTTTGGGGCAACCTTTTATACCGATTTTACTTTATACCACTTTTA 240
Db 181 GAGTAATTCACCTTTGGGGCAACCTTTTATACCGATTTTACTTTATACCACTTTTA 240
QY 241 ACTGATGTTTCACTTTTGGACCGAGTAACTTACCTTTGTTTATTTTGAATAACCGCA 300
Db 241 ACTGATGTTTCACTTTTGGACCGAGTAACTTACCTTTGTTTATTTTGAATAACCGCA 300
QY 301 CTCTCTTCAAGCATATGATGATGACTCGAGTAACTTGTGTTAAAGCGCGCCCTTAAGGA 360
Db 301 CTCTCTTCAAGCATATGATGATGACTCGAGTAACTTGTGTTAAAGCGCGCCCTTAAGGA 360
QY 361 TATCAAGCTTGTACCAAGCGGACACCTTCCACTAGTGTGTTAGTGGATCTGTATC 420
Db 361 TATCAAGCTTGTACCAAGCGGACACCTTCCACTAGTGTGTTAGTGGATCTGTATC 420
QY 421 TCTTCTGAAACATACAGACTAGTATTTATGATCATTTGATCATTTGATGATGATTTCTTGA 480
Db 421 TCTTCTGAAACATACAGACTAGTATTTATGATCATTTGATCATTTGATGATGATTTCTTGA 480
QY 481 AGCGGTTTCACTTTTATACAGAGCTCTTTTGAAGGTGCAATCCATTAATGCGGC 540
Db 481 AGCGGTTTCACTTTTATACAGAGCTCTTTTGAAGGTGCAATCCATTAATGCGGC 540
QY 541 ATAGGTTTCACTTGGCGTATACACTTAACCGTAAACACTTTTGAAGTGGCTGTAT 600
Db 541 ATAGGTTTCACTTGGCGTATACACTTAACCGTAAACACTTTTGAAGTGGCTGTAT 600
QY 601 GCGGCATCTCTTCCGCTACACAGACCTTTTACCAATCTTCTGCTGTGCAACCCACT 660
Db 601 GCGGCATCTCTTCCGCTACACAGACCTTTTACCAATCTTCTGCTGTGCAACCCACT 660
QY 661 GTACGAATAGCATCTACTGCTGTCTGCTGACTTATTTTATATTAAGTAAAAACC 720
Db 661 GTACGAATAGCATCTACTGCTGTCTGCTGACTTATTTTATATTAAGTAAAAACC 720
QY 721 ATTAATATGACAAACACCCCGCTTCACTACCGGTCGAGGAGCGCGCAAGTGGGG 780
Db 721 ATTAATATGACAAACACCCCGCTTCACTACCGGTCGAGGAGCGCGCAAGTGGGG 780
QY 781 TTCAACACGCTCGGACACAGATGCAACCGACCTCCAAAGCCAACTCTGAGCGCGGACC 840
Db 781 TTCAACACGCTCGGACACAGATGCAACCGACCTCCAAAGCCAACTCTGAGCGCGGACC 840
QY 841 GACGACGTAGGACAGGGGTGGCCATTAACGACGTGGCGCATCACTTGTCTTTCCCTT 900
Db 841 GACGACGTAGGACAGGGGTGGCCATTAACGACGTGGCGCATCACTTGTCTTTCCCTT 900
QY 901 TCTCTGCTTCACTTGGCGGCACTGCTGTAGACCCAGGGGATGCTGTGTGAGGAGA 960
Db 901 TCTCTGCTTCACTTGGCGGCACTGCTGTAGACCCAGGGGATGCTGTGTGAGGAGA 960
QY 961 GGTGCGGGGGCCCGATTTTATAGCTTGGCGAGAGAGAGCTTGGCCGCAACCGATCAGA 1020
Db 961 GGTGCGGGGGCCCGATTTTATAGCTTGGCGAGAGAGAGAGCTTGGCCGCAACCGATCAGA 1020
QY 1021 GCTCTGCGCAATATCAAGAAACCAAGTGGGCGCTGCGCGCTAGCCCAACCGCAGAGC 1080
Db 1021 GCTCTGCGCAATATCAAGAAACCAAGTGGGCGCTGCGCGCTAGCCCAACCGCAGAGC 1080
QY 1081 GGGGCTTGTGCGAGCCGTAAGGTGCGGGAAGGGGACCAACCGCTAGGGGGGGCCCATGCTC 1140
Db 1081 GGGGCTTGTGCGAGCCGTAAGGTGCGGGAAGGGGACCAACCGCTAGGGGGGGCCCATGCTC 1140
QY 1141 CACGCGCCAGAGAGAAAAAGAAAGAAAGGCGCGAGATGATG 1183
Db 1141 CACGCGCCAGAGAGAAAAAGAAAGAAAGGCGCGAGATGATG 1183

```

RESULT 2  
US-09-441-340-27/c

```

; Sequence 27, Application US/09441340
; Patent No. 6448476
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/09/441,340
; EARLIER FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 60/108,763
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression
; OTHER INFORMATION: cassette comprising a plant promoter linked to an
; OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
; OTHER INFORMATION: transferase, and a termination sequence
; FEATURE:
; NAME/KEY: Promoter
; LOCATION: (28)..(965)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (966)..(1423)
; FEATURE:
; NAME/KEY: transit peptide
; LOCATION: (1440)..(1667)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1668)..(2099)
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (2114)..(2369)
; US-09-441-340-27

```

Query Match 17.1%; Score 202; DB 3; Length 2378;  
Best Local Similarity 100.0%; Pred. No. 2,5e-89;  
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 163 TTTTATATATGTAAGAAAGATAATTTCACTTTGGGCGCACTTTTATATACGATATTTT 222
Db 202 TTTTATATATGTAAGAAAGATAATTTCACTTTGGGCGCACTTTTATATACGATATTTT 143
QY 223 ACTTATACCACTTTTAACTGATGTTTCACTTTGACAGGTAATCTTACCTTGT 282
Db 142 ACTTATACCACTTTTAACTGATGTTTCACTTTGACAGGTAATCTTACCTTGT 83
QY 283 TATTTTGAATATCCGACTCTTCTCAAGCATATGATGATGATGATGATGATGATGATGAT 342
Db 82 TATTTTGAATATCCGACTCTTCTCAAGCATATGATGATGATGATGATGATGATGATGAT 23
QY 343 ACGGCGCGCGCTTAGGATATC 364
Db 22 ACGGCGCGCGCTTAGGATATC 1

```

RESULT 3  
US-09-068-101-5/c  
Sequence 5, Application US/09068101  
Patent No. 632960  
GENERAL INFORMATION:  
APPLICANT: PLANT GENETIC SYSTEMS N.V.  
TITLE OF INVENTION: Improved Barstar Gene  
FILE REFERENCE: 2121-139P  
CURRENT APPLICATION NUMBER: US/09/068,101  
CURRENT FILING DATE: 1998-08-26  
EARLIER APPLICATION NUMBER: EP 96202446.9  
EARLIER FILING DATE: 1996-09-03  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0



```
SEQ ID NO 5
LENGTH: 4032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: other nucleic
OTHER INFORMATION: acid, "plasmid pmv71"
FEATURE:
NAME/KEY: misc feature
LOCATION: (1995)..(3400)
OTHER INFORMATION: label = PRAC1, "promoter region of rice actin gene"
OTHER INFORMATION: - contains an intron in the leader"
FEATURE:
NAME/KEY: misc feature
LOCATION: (3401)..(3676)
OTHER INFORMATION: label = barstar, "barstar DNA"
FEATURE:
NAME/KEY: misc feature
LOCATION: (3677)..(4003)
OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium
OTHER INFORMATION: T-DNA"
FEATURE:
NAME/KEY: misc feature
LOCATION: (3399)..(3404)
OTHER INFORMATION: label = NcoI, "NcoI recognition site"
FEATURE:
NAME/KEY: misc feature
LOCATION: (4016)..(4021)
OTHER INFORMATION: label = KpnI, "KpnI recognition site"
US-09-970-921-5

Query Match      14.2%; Score 168; DB 3; Length 4032;
Best Local Similarity 100.0%; Pred. No. 1.4e-72;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 222
DB 2166 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 2107

QY 223 ACTTATACCACTTTTAACTGATGTTTCACTTTGACCGATATCTTACCTTTGTT 282
DB 2106 ACTTATACCACTTTTAACTGATGTTTCACTTTGACCGATATCTTACCTTTGTT 2047

QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGA 330
DB 2046 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGA 1999

RESULT 4
US-09-970-921-5/c
Sequence 5, Application US/09970921
Patent No. 6759575
GENERAL INFORMATION:
APPLICANT: Frank Michiels et al.
TITLE OF INVENTION: Improved Barstar Gene
FILE REFERENCE: 2428-0108P
CURRENT APPLICATION NUMBER: US/09/970,921
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 4032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: other nucleic
OTHER INFORMATION: acid, "plasmid pmv71"
NAME/KEY: misc feature
LOCATION: (1999)..(3400)
OTHER INFORMATION: label = PRAC1, "promoter region of rice actin gene"
OTHER INFORMATION: - contains an intron in the leader"
NAME/KEY: misc feature
```

```
LOCATION: (3401)..(3676)
OTHER INFORMATION: label = barstar, "barstar DNA"
NAME/KEY: misc feature
LOCATION: (3677)..(4003)
OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium
OTHER INFORMATION: T-DNA"
NAME/KEY: misc feature
LOCATION: (3399)..(3404)
OTHER INFORMATION: label = NcoI, "NcoI recognition site"
NAME/KEY: misc feature
LOCATION: (4016)..(4021)
OTHER INFORMATION: label = KpnI, "KpnI recognition site"
US-09-970-921-5

Query Match      14.2%; Score 168; DB 4; Length 4032;
Best Local Similarity 100.0%; Pred. No. 1.4e-72;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 222
DB 2166 TTTTAAATAGTGAAGAAAGTAATTTCACTTTGGGCGACCTTTTATACCGATATTT 2107

QY 223 ACTTATACCACTTTTAACTGATGTTTCACTTTGACCGATATCTTACCTTTGTT 282
DB 2106 ACTTATACCACTTTTAACTGATGTTTCACTTTGACCGATATCTTACCTTTGTT 2047

QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGA 330
DB 2046 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGA 1999

RESULT 5
US-09-377-466B-15
Sequence 15, Application US/09377466B
Patent No. 6501009
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/09/377,466B
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 3754
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: expression
OTHER INFORMATION: cassette
NAME/KEY: promoter
LOCATION: (25)..(640)
OTHER INFORMATION: P-CaMV.35S
NAME/KEY: intron
LOCATION: (669)..(1472)
OTHER INFORMATION: I-Zm.Hsp70
NAME/KEY: CDS
LOCATION: (1490)..(3448)
OTHER INFORMATION: Cry3b1 variant v11231
NAME/KEY: terminator
LOCATION: (3475)..(3730)
OTHER INFORMATION: Agrobacterium tumefaciens nos 3' transcription
OTHER INFORMATION: termination and polyadenylation sequence
US-09-377-466B-15

Query Match      12.4%; Score 147; DB 4; Length 3754;
Best Local Similarity 100.0%; Pred. No. 3e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGTTTATGATAGAGTCCCGCAATATATACATTTAATACGC 60
DB 3601 GACGTTATTTATGAGATGGTTTATGATAGAGTCCCGCAATATATACATTTAATACGC 3660
```



QY 61 GATGAAAAAATAATATAGCGCCAACTAGATTAATTATCGCGCGGTGCATCTAT 120  
DB 3022 GATGAAAAAATAATATAGCGCCAACTAGATTAATTATCGCGCGGTGCATCTAT 3081  
QY 121 GTTACTAGATCGGGGATATCCCGGGG 147  
DB 3082 GTTACTAGATCGGGGATATCCCGGGG 3108

RESULT 9  
US-09-186-002-15  
Sequence 15, Application US/09186002B

Patent No. 6489542  
GENERAL INFORMATION:  
APPLICANT: Romano, Charles P.  
TITLE OF INVENTION: Improved Method for Transforming Plants to Express  
FILE REFERENCE: 38-21 (13547) US Pat No. 6489542 09/186,002  
CURRENT FILING DATE: 1998-11-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 10252  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (3660)..(3773)  
OTHER INFORMATION: "n" = g, a, c, or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4355)..(4407)  
OTHER INFORMATION: "n" = g, a, c, or t  
US-09-186-002-15

Query Match 12.4%; Score 147; DB 4; Length 10252;  
Best Local Similarity 100.0%; Pred. No. 2,9e-62;  
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGTTTATGATTAGAGTCCCGCAATTATACATTATTAATACGC 60  
DB 2965 GACGTTATTTATGAGATGGTTTATGATTAGAGTCCCGCAATTATACATTATTAATACGC 3024  
QY 61 GATGAAAAAATAATATAGCGCCAACTAGATTAATTATCGCGCGGTGCATCTAT 120  
DB 3025 GATGAAAAAATAATATAGCGCCAACTAGATTAATTATCGCGCGGTGCATCTAT 3084  
QY 121 GTTACTAGATCGGGGATATCCCGGGG 147  
DB 3085 GTTACTAGATCGGGGATATCCCGGGG 3111

RESULT 10  
US-09-186-002-13  
Sequence 13, Application US/09186002B

Patent No. 6489542  
GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Romano, Charles P.  
TITLE OF INVENTION: Improved Method for Transforming Plants to Express  
FILE REFERENCE: 38-21 (13547) US Pat No. 6489542 09/186,002  
CURRENT FILING DATE: 1998-11-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 10339  
TYPE: DNA  
ORGANISM: Artificial Sequence

FEATURE:  
NAME/KEY: unsure  
LOCATION: (3687)..(3760)  
OTHER INFORMATION: "n" = g, a, c, or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4382)..(4434)  
OTHER INFORMATION: "n" = g, a, c, or t  
US-09-186-002-13

Query Match 12.4%; Score 147; DB 4; Length 10339;  
Best Local Similarity 100.0%; Pred. No. 2,9e-62;  
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGTTTATGATTAGAGTCCCGCAATTATACATTATTAATACGC 60  
DB 3052 GACGTTATTTATGAGATGGTTTATGATTAGAGTCCCGCAATTATACATTATTAATACGC 3111  
QY 61 GATGAAAAAATAATATAGCGCCAACTAGATTAATTATCGCGCGGTGCATCTAT 120  
DB 3112 GATGAAAAAATAATATAGCGCCAACTAGATTAATTATCGCGCGGTGCATCTAT 3171  
QY 121 GTTACTAGATCGGGGATATCCCGGGG 147  
DB 3172 GTTACTAGATCGGGGATATCCCGGGG 3198

RESULT 11  
US-09-441-340-27  
Sequence 27, Application US/09441340

Patent No. 6448426  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
TITLE OF INVENTION: Phosphonate Metabolizing Plants  
FILE REFERENCE: 38-21 (15103)  
CURRENT FILING DATE: 1999-11-16  
EARLIER APPLICATION NUMBER: 60/108,763  
EARLIER FILING DATE: 1998-11-17  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 27  
LENGTH: 2378  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (966)..(1423)  
OTHER INFORMATION: Description of Artificial Sequence: expression  
OTHER INFORMATION: cassette comprising a plant promoter linked to an  
OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl  
OTHER INFORMATION: transferase, and a termination sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (128)..(965)  
FEATURE:  
NAME/KEY: intron  
LOCATION: (966)..(1423)  
FEATURE:  
NAME/KEY: transic peptide  
LOCATION: (1440)..(1667)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1668)..(2099)  
FEATURE:  
NAME/KEY: terminator  
LOCATION: (2114)..(2369)  
US-09-441-340-27

Query Match 11.8%; Score 140; DB 3; Length 2378;  
Best Local Similarity 100.0%; Pred. No. 8,3e-59;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGTTTATGATTAGAGTCCCGCAATTATACATTATTAATACGC 60

Db 2239 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 2239  
QY 61 GATAGAAAACAAATATAGCGCGCAACTAGATTAATATTCGCGCGGTGTCTAT 120  
Db 2299 GATAGAAAACAAATATAGCGCGCAACTAGATTAATATTCGCGCGGTGTCTAT 2358  
QY 121 GTTACTAGATCGGGGATATC 140  
Db 2359 GTTACTAGATCGGGGATATC 2378

RESULT 12  
US-09-441-340-31  
; Sequence 31, Application US/09441340  
; Patent No. 6448476  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; TITLE OF INVENTION: Phosphonate Metabolizing Plants  
; FILE REFERENCE: 38-21(15303)  
; CURRENT APPLICATION NUMBER: US/09/441,340  
; EARLIER FILING DATE: 1999-11-16  
; EARLIER APPLICATION NUMBER: 60/108,763  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 2436  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: monocot  
; OTHER INFORMATION: expression cassette comprising plant operable  
; OTHER INFORMATION: promoter linked to an intron, a sequence coding  
; OTHER INFORMATION: for an AMPA acetyltransferase, and a termination  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (26)..(640)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (670)..(1473)  
; FEATURE:  
; NAME/KEY: transic\_peptide  
; LOCATION: (1498)..(1725)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1726)..(2157)  
; FEATURE:  
; NAME/KEY: terminator  
; LOCATION: (2172)..(2427)  
; US-09-441-340-31

Query Match 11.8%; Score 140; DB 3; Length 2436;  
Best Local Similarity 100.0%; Pred. No. 8.5e-59;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 60  
Db 2297 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 2356  
QY 61 GATAGAAAACAAATATAGCGCGCAACTAGATTAATATTCGCGCGGTGTCTAT 120  
Db 2357 GATAGAAAACAAATATAGCGCGCAACTAGATTAATATTCGCGCGGTGTCTAT 2416  
QY 121 GTTACTAGATCGGGGATATC 140  
Db 2417 GTTACTAGATCGGGGATATC 2436

RESULT 13  
US-09-441-340-23  
; Sequence 23, Application US/09441340  
; Patent No. 6448476

; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; TITLE OF INVENTION: Phosphonate Metabolizing Plants  
; FILE REFERENCE: 38-21(15303)  
; CURRENT APPLICATION NUMBER: US/09/441,340  
; EARLIER FILING DATE: 1999-11-16  
; EARLIER APPLICATION NUMBER: 60/108,763  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 1630  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression  
; OTHER INFORMATION: cassette comprising plant operable promoter linked  
; OTHER INFORMATION: to a coding sequence encoding an AMPA  
; OTHER INFORMATION: acetyltransferase linked to a transcription  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (33)..(605)  
; FEATURE:  
; NAME/KEY: transic\_peptide  
; LOCATION: (627)..(892)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (893)..(1324)  
; FEATURE:  
; NAME/KEY: terminator  
; LOCATION: (1350)..(1605)  
; US-09-441-340-23

Query Match 11.6%; Score 137; DB 3; Length 1630;  
Best Local Similarity 100.0%; Pred. No. 2.6e-57;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 60  
Db 1476 GACGTTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTAAATACGC 1535  
QY 61 GATAGAAAACAAATATAGCGCGCAACTAGATTAATATTCGCGCGGTGTCTAT 120  
Db 1536 GATAGAAAACAAATATAGCGCGCAACTAGATTAATATTCGCGCGGTGTCTAT 1595  
QY 121 GTTACTAGATCGGGAT 137  
Db 1596 GTTACTAGATCGGGAT 1612

RESULT 14  
US-09-098-2198-5  
; Sequence 5, Application US/090982198  
; Patent No. 6441277  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard  
; APPLICANT: Cheikh, No. 6441277dine  
; APPLICANT: Kishore, Ganesh  
; TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Arnold White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TX  
; COUNTRY: US  
; ZIP: 77210-4433  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/098,219B  
 FILING DATE:  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/049,995  
 FILING DATE: 17-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kammerer, Patricia A.  
 REGISTRATION NUMBER: 29,775  
 REFERENCE/DOCKET NUMBER: MOBT-086  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 713-787-1400  
 TELEFAX: 713-787-1440  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 10846 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-09-098-219B-5

Query Match 11.6%; Score 137; DB 3; Length 10846;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-57;  
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGATCCCGCAATTATATACATTATTAATACGC 60  
 DB 4489 GACGTTATTATGAGATGGGTTTATGATTAGATCCCGCAATTATATACATTATTAATACGC 4548  
 QY 61 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120  
 DB 4549 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 4608  
 QY 121 GTTACTAGATCGGGGAT 137  
 DB 4609 GTTACTAGATCGGGGAT 4625

RESULT 15  
 US-10-164-204-5  
 Sequence 5, Application US/10164204  
 Patent No. 6663906  
 GENERAL INFORMATION:  
 APPLICANT: Cheikh, No. 6663906dine  
 APPLICANT: Kishore, Ganesh  
 TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate Aldolase in Transgenic P  
 FILE REFERENCE: 11899, 0086, DUS02 (MOBT:086-2)  
 CURRENT APPLICATION NUMBER: US/10/164,204  
 CURRENT FILING DATE: 2002-06-06  
 PRIOR APPLICATION NUMBER: 09/098,219  
 PRIOR FILING DATE: 1998-06-16  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: Patentin version 3.1  
 SEQ ID NO 5  
 LENGTH: 10846  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: P-FMV/CTPI/Ida/NOS3,  
 US-10-164-204-5

Query Match 11.6%; Score 137; DB 4; Length 10846;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-57;  
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTATGAGATGGGTTTATGATTAGATCCCGCAATTATATACATTATTAATACGC 60  
 DB 4489 GACGTTATTATGAGATGGGTTTATGATTAGATCCCGCAATTATATACATTATTAATACGC 4548  
 QY 61 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 120

DB 4549 GATGAAAAAATAATATAGCGCGCAACTAGATTAATTATCGCGCGGTGCATCTAT 4608  
 QY 121 GTTACTAGATCGGGGAT 137  
 DB 4609 GTTACTAGATCGGGGAT 4625

Search completed: February 10, 2005, 08:49:47  
 Job time : 365.533 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 121.68 Seconds  
(without alignments)  
5630.828 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18 tgcgtctcgtcgtcgtc 18

Sequence: 1 tgcgtctcgtcgtcgtc 18

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gest1.\*  
9: gb\_gest2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 18    | 100.0       | 354    | BF450993 | BF450993 uz73a01.y |
| 2          | 18    | 100.0       | 474    | AA103332 | AA103332 mo24d04.x |
| 3          | 18    | 100.0       | 517    | BM507605 | BM507605 ih42b06.y |
| 4          | 18    | 100.0       | 518    | BM626708 | BM626708 uu10e06.y |
| 5          | 18    | 100.0       | 522    | CF916734 | CF916734 B0999R07- |
| 6          | 18    | 100.0       | 529    | CA557138 | CA557138 K0222G09- |
| 7          | 18    | 100.0       | 534    | BE632036 | BE632036 uu10e06.x |
| 8          | 18    | 100.0       | 540    | BI966323 | BI966323 id50h12.x |
| 9          | 18    | 100.0       | 544    | BU605257 | BU605257 mah89b07- |
| 10         | 18    | 100.0       | 557    | CF171411 | CF171411 B0842B01- |
| 11         | 18    | 100.0       | 576    | AM559055 | AM559055 L0303E07- |
| 12         | 18    | 100.0       | 598    | BM219770 | BM219770 C0329G10- |
| 13         | 18    | 100.0       | 603    | CF739008 | CF739008 UI-M-HD0  |
| 14         | 18    | 100.0       | 605    | CF916509 | CF916509 B0996C03- |
| 15         | 18    | 100.0       | 609    | BM237875 | BM237875 K0508E08- |
| 16         | 18    | 100.0       | 610    | CA888455 | CA888455 B0145C12- |
| 17         | 18    | 100.0       | 617    | BM601036 | BM601036 BM601036  |
| 18         | 18    | 100.0       | 617    | BM507714 | BM507714 ih42b06.x |
| 19         | 18    | 100.0       | 629    | BI650935 | BI650935 603297263 |
| 20         | 18    | 100.0       | 629    | BU702841 | BU702841 UI-M-FC0  |
| 21         | 18    | 100.0       | 630    | BM076391 | BM076391 H3159C09- |
| 22         | 18    | 100.0       | 630    | BM233685 | BM233685 K0340B05- |
| 23         | 18    | 100.0       | 668    | CNS31891 | CNS31891 UI-M-HQ0- |
| 24         | 18    | 100.0       | 709    | BQ770022 | BQ770022 UI-M-F10- |

|    |    |       |      |          |                     |
|----|----|-------|------|----------|---------------------|
| 25 | 18 | 100.0 | 742  | BI853956 | BI853956 603381039  |
| 26 | 18 | 100.0 | 812  | BI872911 | BI872911 603398262  |
| 27 | 18 | 100.0 | 907  | BF179971 | BF179971 601806316  |
| 28 | 18 | 100.0 | 997  | BQ713665 | BQ713665 AGENCOURT  |
| 29 | 18 | 100.0 | 2237 | AK036132 | AK036132 Mus muscu  |
| 30 | 17 | 94.4  | 604  | BM236573 | BM236573 K0437G06-  |
| 31 | 17 | 94.4  | 655  | AZ572271 | AZ572271 304PvB11   |
| 32 | 16 | 88.9  | 182  | BQ321075 | BQ321075 CM0-CT030  |
| 33 | 16 | 88.9  | 182  | BQ321079 | BQ321079 CM0-CT030  |
| 34 | 16 | 88.9  | 297  | AQ421868 | AQ421868 RPEC1-11-1 |
| 35 | 16 | 88.9  | 433  | AQ298483 | AQ298483 HS 2229 B  |
| 36 | 16 | 88.9  | 503  | BF289471 | BF289471 EST45062   |
| 37 | 16 | 88.9  | 537  | CL409705 | CL409705 RPEC14_41  |
| 38 | 16 | 88.9  | 646  | BM197672 | BM197672 Danio rer  |
| 39 | 16 | 88.9  | 653  | BM122080 | BM122080 Danio rer  |
| 40 | 16 | 88.9  | 672  | AZ835346 | AZ835346 2M0129H05  |
| 41 | 16 | 88.9  | 761  | BM142562 | BM142562 Danio rer  |
| 42 | 16 | 88.9  | 782  | CA384083 | CA384083 BX913369   |
| 43 | 16 | 88.9  | 796  | CNS04FOG | CNS04FOG Tetraodon  |
| 44 | 16 | 88.9  | 796  | CNS04FOG | CNS04FOG Tetraodon  |
| 45 | 16 | 88.9  | 811  | B2175244 | B2175244 CH230-455  |

#### ALIGNMENTS

RESULT 1  
BF450993 354 bp mRNA linear EST 29-DEC-2000  
uz73a01.y1 NCI\_CGAP\_lu29 Mus musculus cDNA clone IMAGE:3674664 5',  
DEFINITION  
mRNA sequence.  
BF450993 GI:11517162

ACCESSION  
BF450993  
VERSION  
BF450993.1 GI:11517162  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
1 (bases 1 to 354)  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL  
COMMENT  
Other ESTs: uz73a01.x1  
Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
image.llnl.gov/image/html/iresources.shtml

MG1:1435432  
Seq primer: -40RP from Gibco  
High quality sequence stop: 340.  
Location/Qualifiers

#### FEATURES

source  
1..354  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3674664"  
/issue\_type="spontaneous tumor, metastatic to mammary."  
Stem cell origin."  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_lu29"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
 |||||  
 40 TGCTGTTCTGCTGACTTT 57

RESULT 2 474 bp mRNA linear EST 29-OCT-1996  
 AA103332  
 LOCUS AA103332  
 DEFINITION mo2404.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus  
 CDNA clone IMAGE:554503 5', mRNA sequence.

ACCESSION AA103332 GI:1649616  
 VERSION AA103332.1  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 474)  
 Author: M. Hillier, L. Allen, M. Bowles, M. Dietrich, N. Dubuque, T. Geisel, S. Kucaba, T. Lacy, M. Le, M. Martin, J. Morris, M. Schellenberg, K. Steptoe, M. Tan, F. Underwood, K. Moore, B. Theising, B. Wylie, T. Lennon, G. Soares, B. Wilson, R. and Waterston, R.  
 The WashU-HMNI Mouse EST Project  
 Unpublished (1996)  
 Contact: Marra M/Mouse EST Project  
 WashU-HMNI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:335295  
 Seq primer: -28M13 rev1 from Amerisham  
 High quality sequence stop: 430.  
 Location/Qualifiers

FEATURES  
 source  
 1..474  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:554503"  
 /issue\_type="embryo"  
 /dev\_stage="13.5dpc embryos"  
 /lab\_host="DH10B"  
 /clone\_id="Life Tech mouse embryo 13 5dpc 10666014"  
 /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site: 1; Salt; Site 2: NciI; Cloned unidirectionally. Primer: Oligo dt. 13.5dpc embryos. pCMV-SPORT2 vector."

ORIGIN  
 Query Match 100.0%; Score 18; DB 1; Length 474;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
 |||||  
 337 TGCTGTTCTGCTGACTTT 354

RESULT 3 517 bp mRNA linear EST 12-MAR-2002  
 BM507605  
 LOCUS BM507605  
 DEFINITION In42506.y1 Melton Mouse E16 5 Pancreas Library 2 M1652 Mus musculus  
 CDNA clone IMAGE:5679706 5', mRNA sequence.

ACCESSION BM507605

VERSION BM507605.1 GI:18678748  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 517)  
 Author: Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blais, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Williams, T., Jackson, Y., McCann, R., Cole, K., Tsagaris, R., Williams, T., Jackson, Y., and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biochem.harvard.edu  
 Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brownjfas.harvard.edu)  
 MGI:1957584 This sequence now available from the IMAGE consortium, see clone orders contact: info@image.llnl.gov  
 Seq primer: -40BP from Gibco  
 High quality sequence stop: 432.  
 Location/Qualifiers

FEATURES  
 source  
 1..517  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5679706"  
 /sex="Both"  
 /issue\_type="Total pancreas"  
 /dev\_stage="Embryonic day 16.5"  
 /lab\_host="TOP10"  
 /clone\_id="Melton Mouse E16 5 Pancreas Library 2 M1652"  
 /note="Organ: Pancreas; Vector: pBluescript II SK; Site 1: NciI; Site 2: SalI; Library constructed using Superscript Plasmid Library kit (Life Technologies). CDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.06Kb. Primary library, unamplified."

ORIGIN  
 Query Match 100.0%; Score 18; DB 4; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
 |||||  
 430 TGCTGTTCTGCTGACTTT 447

RESULT 4 518 bp mRNA linear EST 24-AUG-2000  
 BE626708  
 LOCUS BE626708  
 DEFINITION un10e06.y2 Soares mammary\_gland NMIMG Mus musculus CDNA clone  
 IMAGE:3371554 5' similar to SW:KRU1\_DROME P08155 KRUPEL  
 HOMOLOGOUS PROTEIN; , mRNA sequence.

ACCESSION BE626708  
 VERSION BE626708.1 GI:9907124  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.



REFERENCE 1 (bases 1 to 518)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 TITLE Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapdb-rc@mail.nih.gov](mailto:cgapdb-rc@mail.nih.gov)  
 This clone is available royalty-free through LML; contact the  
 IMAGE Consortium ([info@image.lml.gov](mailto:info@image.lml.gov)) for further information.  
 MGI:1081158  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 456.

## FEATURES

source

Location/Qualifiers

1..518  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3371554"  
 /sex="female (lactating)"  
 /tissue\_type="mammary gland"  
 /lab\_host="DH10B"  
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified  
 polylinker; 1st strand cDNA was prepared from mammary  
 gland tissue from a lactating female, and was then primed  
 with a Not I - oligo (dT) primer. Double-stranded cDNA was  
 ligated to Eco RI adaptors (Pharmacia), digested with Not  
 I and cloned into the Not I and Eco RI sites of the  
 modified pRT3D vector. Library is normalized. Library  
 was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGCTGTTTCGCTGACTTT 18  
 ||||||||||||||||  
 Db 493 TGCTGTTTCGCTGACTTT 510

RESULT 5  
 CF916734 522 bp mRNA linear EST 05-NOV-2003  
 LOCUS B0999F07-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus  
 DEFINITION musculus cDNA clone NIA:B0999F07 IMAGE:30481794 5', mRNA sequence.  
 CF916734  
 ACCESSION CF916734.1 GI:38187936  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 522)  
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
 Construction of long-transcript enriched cDNA libraries from  
 submicrogram amounts of total RNAs by a universal PCR amplification  
 method  
 Genome Res. 11 (9), 1553-1558 (2001)  
 JOURNAL MEDLINE  
 MEDLINE 21429098  
 PUBMED 11544199  
 COMMENT Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Caswell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: [cdna@gsun.grc.nia.nih.gov](mailto:cdna@gsun.grc.nia.nih.gov)  
 Plate: B0999 row: F column: 07  
 Seq primer: M13 Reverse  
 High quality sequence stop: 522  
 POLYA=No.

## FEATURES

source

Location/Qualifiers

1..522  
 /organism="Mus musculus"

/mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:B0999F07-5"  
 /db\_xref="taxon:10090"  
 /clone="NIA:B0999F07 IMAGE:30481794"  
 /dev\_stage="Unfertilized Egg"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse Unfertilized Egg cDNA Library (Long  
 1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;  
 Site 2: NotI; Mouse cDNA project by the Laboratory of  
 Genetics, National Institute on Aging (NIA), Intramural  
 Research Program, NIH (<http://gsun.grc.nia.nih.gov/cDNA>).  
 This is a long-transcript enriched cDNA library (Ref.  
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total  
 RNAs were extracted from a pool of 1488 unfertilized eggs.  
 Double-stranded cDNAs were synthesized with an Oligo(dT)  
 primer (Invitrogen):  
 5'-pGACTGTTCTAGATCGAGCGGCCCTTTTTTTTTTT-3'.  
 treated with T4 DNA polymerase, and purified by  
 ethanol-precipitation. The cDNAs were ligated to  
 Lene-linker L-SalI, purified by phenol/chloroform, and  
 separated from free linkers by Centricon 100. Then, the  
 cDNAs were amplified by long-range high fidelity PCR using  
 Ex Taq polymerase (Takara) with a primer Sal14-5. The  
 products were purified by phenol/chloroform and Centricon  
 100. The cDNAs were digested with SalI and NotI enzymes  
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid  
 vector. The DH10B E. coli host was transformed with the  
 ligation mixture by the standard chemical method. The  
 average insert size is about 2.5 kb. The library was  
 constructed by Yulan Piao."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGCTGTTTCGCTGACTTT 18  
 ||||||||||||||||  
 Db 183 TGCTGTTTCGCTGACTTT 200

RESULT 6  
 CAS57138 529 bp mRNA linear EST 19-NOV-2002  
 LOCUS K0222G09-5N NIA Mouse Unfertilized Egg cDNA Library (Long) Mus  
 DEFINITION musculus cDNA clone NIA:K0222G09 IMAGE:30046832 5', mRNA sequence.  
 CAS57138  
 ACCESSION CAS57138.1 GI:25101573  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 529)  
 Piao, Y., Kargu, G.J., Dudekula, D.B., Qian, Y., Luo, A., Stagg, C.A.,  
 Martin, P., Aiba, K., Tanaka, T. and Ko, M.S.H.  
 Systematic Analyses of NIA Mouse Unfertilized Egg cDNA Library  
 (Long)  
 Unpublished (2001)  
 Other ESTs: K0222G09-3  
 COMMENT Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Caswell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: [cdna@gsun.grc.nia.nih.gov](mailto:cdna@gsun.grc.nia.nih.gov)  
 Plate: K0222 row: G column: 09  
 Seq primer: M13 Reverse  
 High quality sequence stop: 529  
 POLYA=No.

## FEATURES

source

Location/Qualifiers

1..529

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="niaEST:K0222G09-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0222G09 IMAGE:30046832"  
/issue\_type="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_1lb="NIA Mouse Unfertilized Egg cDNA Library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://igun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001) [PMID: 11541991]). Total RNA were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen):  
5'-GCACTAGTCTGATCGGAGCGCCGCTTTTCTTTTCTTTT-3',  
treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 529;  
Best Local Similarity, 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 211 TGCTGTTCTGCTGACTTT 228

RESULT 7  
BE632036/c 534 bp mRNA linear EST 25-AUG-2000  
LOCUS BE632036  
DEFINITION IMAGE:3371554 3', mRNA sequence.  
ACCESSION BE632036  
VERSION BE632036  
KEYWORDS GI:9914724  
SOURCE EST.  
ORGANISM Mus musculus (house mouse)  
MUS musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 534)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Other ESTs: un10e06.y2  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
This clone is available royalty-free through LIND; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:1081158  
High quality sequence stop: 469.

FEATURES  
source  
1..534  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"

/clone="IMAGE:3371554"  
/sex="female (lactating)"  
/issue\_type="mammary gland"  
/lab\_host="DH10B"  
/clone\_1lb="Soares mammary gland NMLMG"  
/note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptor (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library is normalized. Library was constructed by Bento Soares and W. Fatima Bonalio."

## ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 534;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 506 TGCTGTTCTGCTGACTTT 489

RESULT 8  
B1966323/c 540 bp mRNA linear EST 12-MAR-2002  
LOCUS B1966323  
DEFINITION i450h12.x1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus  
musculus cDNA clone IMAGE:5666926 3', mRNA sequence.  
ACCESSION B1966323  
VERSION B1966323.1 GI:16340728  
KEYWORDS EST.  
SOURCE  
ORGANISM Mus musculus (house mouse)  
MUS musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 540)  
Melton, D., Brown, J., Keny, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scarce, M., Breasted, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blaisdell, A., Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, T., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: [dmelton@biohp.harvard.edu](mailto:dmelton@biohp.harvard.edu)  
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown ([brown@fas.harvard.edu](mailto:brown@fas.harvard.edu))  
MGI:1953252 This sequence now available from the IMAGE consortium, for clone orders contact: [info@image.llnl.gov](mailto:info@image.llnl.gov)  
High quality sequence stop: 449.

## FEATURES

source  
1..540  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="ICR"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5666926"  
/sex="Both for embryonic & newborn, male for adult and adult islet"  
/dev stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"  
/lab\_host="DH10B"  
/clone\_1lb="Melton Normalized Mixed Mouse Pancreas 1"

NI-MMS1"  
 /note="Vector: pSPORT1; Site 1: Not 1; Site 2: Sal I; Five  
 libraries representing E10.5/12.5 pancreatic bud, E16.5  
 pancreas, newborn pancreas, adult pancreas, and adult  
 islets of Langerhans were separately constructed using  
 Superscript Plasmid Library kit (Life Technologies). cDNA  
 was made by oligo-dT priming and size-selected by column  
 fractionation. Libraries were amplified once on solid  
 support and plasmid DNA from each library was prepared  
 and mixed in equal amounts. The mixed library DNA was  
 normalized by method #4 from Bonaldo, Lennon, and Soares  
 1996 Genome Research 6:791-806; 0.5 microgram  
 single-stranded mixed library plasmid DNA was mixed with  
 5 micrograms PCR product representing mixed library  
 inserts and hybridized to an EcoT of 6. Single-stranded  
 (unhybridized) plasmids were isolated by hydroxyapatite  
 chromatography and used to make this library."

## ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 540;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
 |||||  
 Db 504 TGCTGTTCTGCTGACTTT 487

RESULT 9  
 LOCUS BU605257 544 bp mRNA linear EST 20-SEP-2002  
 DEFINITION mah89b07.y1 McCarrey Eddy 18 day leptotene and zygotene  
 spermatocytes Mus musculus cDNA clone IMAGE:6367453 5', mRNA  
 sequence.  
 BU605257  
 BU605257.1 GI:23261437

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS  
 McCarrey, J., Eddy, M., Marra, M., Hillier, L., Clifton, S., Pape, D.,  
 Martin, J., Wylie, T., Dante, M., Bowers, Y., Theising, B., Gibbons, M.,  
 Rutter, E., Tsagaris, V., Ronko, I., Maguire, L., Kennedy, S.,  
 Bennett, J., Waterston, R. and Wilson, R.

TITLE  
 JOURNAL  
 COMMENT  
 Unpublished (2002)  
 Contact: McCarrey/Eddy NIH Mouse  
 NIH Mouse  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

Library constructed and donated by J. McCarrey, Ph.D. (Southwest  
 Foundation for Biomedical Research, Dept. of Genetics) - excision  
 done by E. M. Eddy, Ph.D. (National Institutes of Health, National  
 Institute of Environmental Health Sciences).  
 MGI:2044885  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 1..544

FEATURES  
 source  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6367453"  
 /sex="male"  
 /issue\_type="18-day leptotene and zygotene spermatocytes"  
 /lab\_host="VDH08 (phage-resistant)"  
 /clone\_lib="McCarrey Eddy 18 day leptotene and zygotene  
 spermatocytes"

/note="Organ: testis; Vector: pBluescript SK+  
 (Stratagene); Site 1: EcoRI; Site 2: XhoI; cDNA oligo  
 dt-primed [5'-(GA)10-ACGTGCTGCGATTTTCTTTT-3'] and  
 directionally cloned using 5' linkers 5'-AATCGGACGAG-3'  
 and 5'-CTCGGCG-3'. Size selection of >400bp material  
 gives average insert size ranging from 1-2 kb. Library was  
 mass excised (from lambda-UniZAP-XR) and resulting  
 single-stranded phagmids were prepped and transformed into  
 DH10B. Library constructed and donated by J. McCarrey,  
 Ph.D. (Southwest Foundation for Biomedical Research, Dept.  
 of Genetics); excision done by E. M. Eddy, Ph.D. (National  
 Institutes of Health, National Institute of Environmental  
 Health Sciences)."

## ORIGIN

Query Match 100.0%; Score 18; DB 5; Length 544;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
 |||||  
 Db 341 TGCTGTTCTGCTGACTTT 358

RESULT 10  
 LOCUS CP171411 557 bp mRNA linear EST 25-JUL-2003  
 DEFINITION B0842B01-5 NIA Mouse Newborn Kidney cDNA library (long 1) Mus  
 musculus cDNA clone NIA:B0842B01 IMAGE:30471468 5', mRNA sequence.  
 CP171411  
 CP171411.1 GI:33280960

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS  
 TITLE  
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
 1 (bases 1 to 557)  
 Construction of long-transcript enriched cDNA libraries from  
 submicrogram amounts of total RNAs by a universal PCR amplification  
 method  
 Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 Contact: Dawood B. Dudekula  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@lgsun.grc.nia.nih.gov  
 Plate: B0842 row: B column: 01  
 Seq primer: M13 Reverse  
 High quality sequence stop: 557  
 POLYA=No.

FEATURES  
 source  
 Location/Qualifiers  
 1..557

FEATURES  
 source  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="NIA:B0842B01-5"  
 /dev\_stage="Newborn Kidney"  
 /lab\_host="VDH08"  
 /clone\_lib="NIA Mouse Newborn Kidney cDNA library (long  
 1)"  
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;  
 Site 2: NotI; Mouse cDNA project by the Laboratory of  
 Genetics, National Institute on Aging (NIA), Intramural  
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).  
 In brief, double-stranded cDNAs were synthesized with an  
 Oligo(dT) primer (Invitrogen):  
 5'-pGACTAGTTCAGATCGGACGCGCCCTTTTCTTTT-3' from  
 26 ug of total RNA, treated with T4 DNA polymerase, and

## ORIGIN

purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-5. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

Query Match 100.0%; Score 18; DB 7; Length 557;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18  
|||||

Db 447 TGCTGTTCTGCTGACTTT 464

RESULT 11  
AM559055/c 576 bp mRNA linear EST 31-AUG-2000

LOCUS L0303E07-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA

DEFINITION clone L0303E07 3', mRNA sequence.

ACCESSION AM559055

VERSION AM559055.1 GI:7204484

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 576); Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X.,

Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagata, R., Doi, R.,

Wood, W.H., III, Becker, K.G., and Ko, M.S.H.

Genome-wide expression profiling of mid-gestation placenta and

embryo using a 15,000 mouse developmental cDNA microarray

Proc. Natl. Acad. Sci. U.S.A. 97 (Feb. 9, 2000) 9127-9132 (2000)

20381348

10922068

COMMENT Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: L0303 row: E column: 07

Seq primer: -21M13 Forward

High quality sequence stop: 576

POLYA=yes.

Location/Qualifiers

1..576

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="L0303E07"

/sex="Female"

/dev\_stage="Newborn Ovary"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse Newborn Ovary cDNA Library"

/note="Vector: pSPORT1 (Gibco/BRL life technology);

Site 1: SalI; Site 2: NotI; Total RNAs were extracted from

7 Newborn Ovary. The double-stranded cDNA was synthesized

RESULT 12

BM219770/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 598);

Piao, Y., Ko, N.T., Lim, M.K., and Ko, M.S.H.

Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

21429098

11544199

COMMENT On Dec 14, 2001 this sequence version replaced gi:17779692.

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: C0929 row: G column: 10

Seq primer: -21M13 Forward

High quality sequence stop: 598

POLYA=yes.

Location/Qualifiers

1..598

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="NIA:C0929G10 IMAGE:30037425"

/sex="Male"

/tissue\_type="Male genital ridge/meonephros"

/dev\_stage="12.5-dpc"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse 12.5-dpc Male Genital

ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Yulan Piao."

Query Match 100.0%; Score 18; DB 2; Length 576;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18  
|||||

Db 520 TGCTGTTCTGCTGACTTT 503

BM219770 598 bp mRNA linear EST 07-JUN-2003

C0929G10-3 NIA Mouse 12.5-dpc Male Genital Ridge/Meonephros cDNA

Library (long) Mus musculus cDNA clone NIA:C0929G10 IMAGE:30037425

3', mRNA sequence.

BM219770.2 GI:31478468

EST.

SOURCE Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 598);

Piao, Y., Ko, N.T., Lim, M.K., and Ko, M.S.H.

Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

21429098

11544199

COMMENT On Dec 14, 2001 this sequence version replaced gi:17779692.

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: C0929 row: G column: 10

Seq primer: -21M13 Forward

High quality sequence stop: 598

POLYA=yes.

Location/Qualifiers

1..598

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="NIA:C0929G10 IMAGE:30037425"

/sex="Male"

/tissue\_type="Male genital ridge/meonephros"

/dev\_stage="12.5-dpc"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse 12.5-dpc Male Genital

Ridge/Meonephros cDNA Library (long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is

a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pgactggtctgacatcgccgagccgccctttttttttttt-3'] from 1.8 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-4-S. The products were purified by phenol/chloroform and Gentricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 598;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTTCTGACTTT 18  
|||||  
Db 504 TGCTGTTTCTGACTTT 487

RESULT 13  
LOCUS CF739008 603 bp mRNA linear EST 10-OCT-2003  
DEFINITION U1-M-HD0-cku-p-16-0-U1.r1 NIH BMAP\_HD0 Mus musculus cDNA clone  
IMAGE:30610839 5', mRNA sequence.  
CF739008  
ACCESSION CF739008.1 GI:37635345  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 603)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. James Lin University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mouse1.html  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

FEATURES  
source Location/Qualifiers  
1..603  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30610839"  
/tissue\_type="whole eye"  
/dev\_stage="embryo 12.5, 13.5, 14.5 dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_1id="NIH-BMAP\_HD0"  
/note="Organ: Eye; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTTATTAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH)."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 603;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTTCTGACTTT 18  
|||||  
Db 117 TGCTGTTTCTGACTTT 134

RESULT 14  
LOCUS CF916509 605 bp mRNA linear EST 05-NOV-2003  
DEFINITION B0996C03-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus musculus cDNA clone NIA:B0996C03 IMAGE:30481466 5', mRNA sequence.  
CF916509  
ACCESSION CF916509.1 GI:38187711  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 605)  
AUTHORS Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.  
TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method  
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)  
MEDLINE 21429098  
PUBMED 11544199  
COMMENT Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@genetics.nia.nih.gov  
Plate: B0996 row: C column: 03  
Seq primer: M13 Reverse  
High quality sequence stop: 605  
POLYA=No.

FEATURES  
source Location/Qualifiers  
1..605  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="NIA:B0996C03 IMAGE:30481466"  
/dev\_stage="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_1id="NIA Mouse Unfertilized Egg cDNA Library (Long 1)"  
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igen.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNA were extracted from a pool of 148 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen: 5'-pgactggtctgacatcgccgagccgccctttttttttt-3'), treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker L1-SalI, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

## ORIGIN

Query Match 100.0%; Score 18; DB 7; Length 609;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 211 TGCTGTTCTGCTGACTTT 228

RESULT 15 609 bp mRNA linear EST 07-JUN-2003  
BM237875/c K0508E08-3 NIA Mouse Hematopoietic Stem Cell (Lin-/C-Kit+/Sca-1+)  
LOCUS CDNA Library (Long) Mus musculus CDNA clone NIA:K0508E08  
DEFINITION IMAGE:30064663 3', mRNA sequence.

ACCESSION BM237875 GI:31494430  
VERSION BM237875.2  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 609)  
AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.  
TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)  
MEDLINE 21429098  
PUBMED 11544199  
COMMENT On Dec 17, 2001 this sequence version replaced gi:17873145.  
Other ESTs: K0508E08-5N

Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0508 row: E column: 08  
Seq primer: -21M13 Forward  
High quality sequence stop: 609  
POLYAs=Yes.

## FEATURES

## Source

1. 609  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6Ncr"  
/db\_xref="niaest:K0508E08-3"  
/db\_xref="taxon:10090"  
/clone="NIA:K0508E08 IMAGE:30064663"  
/issue\_type="Hematopoietic Stem Cell  
(Lin-/C-Kit+/Sca-1+)"  
/dev\_stage="Age approx.10 weeks old"  
/lab\_host="DH10B"  
/clone\_id="NIA Mouse Hematopoietic Stem Cell  
(Lin-/C-Kit+/Sca-1+)"  
/note="Vector: pSPORT (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an Oligo(dT) primer (Invitrogen):  
5'-pGACTACTTCTAGATCGCGCGCCGCCCTTTTCTTTT-3' from  
4.8 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to lone-linker L1-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.7 kb. The library was constructed  
by Yulan Piao (NIA)."

## ORIGIN

Query Match 100.0%; Score 18; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 504 TGCTGTTCTGCTGACTTT 487

Search completed: February 10, 2005, 17:01:59  
Job time : 124.68 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 3071.76 Seconds  
(without alignments)  
7855.666 Million cell updates/sec

Title: US-10-790-430-7

Perfect score: 498  
Sequence: 1 aatcgatcccaaatcgcgcac.....aaaagctgcccagaatgaa 498

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:\*  
1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_str:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description          |
|------------|-------|-------------|--------|-----------|----------------------|
| 1          | 498   | 100.0       | 498    | AX342368  | AX342368 Sequence    |
| 2          | 118   | 23.7        | 1183   | AX342369  | AX342369 Sequence    |
| 3          | 101   | 20.3        | 2378   | AR229555  | AR229555 Sequence    |
| 4          | 70    | 14.1        | 1266   | S44221    | S44221 Actin         |
| 5          | 70    | 14.1        | 1404   | IS0115    | IS0115 Sequence      |
| 6          | 70    | 14.1        | 1501   | AX044094  | AX044094 Sequence    |
| 7          | 70    | 14.1        | 1501   | AX044144  | AX044144 Sequence    |
| 8          | 70    | 14.1        | 2199   | IS0113    | IS0113 Sequence      |
| 9          | 70    | 14.1        | 4895   | AY452735  | AY452735 Reporter    |
| 10         | 70    | 14.1        | 5643   | IS0112    | IS0112 Sequence      |
| 11         | 70    | 14.1        | 9359   | AX384394  | AX384394 Sequence    |
| 12         | 70    | 14.1        | 9359   | AX473364  | AX473364 Sequence    |
| 13         | 70    | 14.1        | 147344 | AC091532  | AC091532 Oryza sat   |
| 14         | 68    | 13.7        | 1392   | IS0114    | IS0114 Sequence      |
| 15         | 68    | 13.7        | 4412   | EVCOR112N | EVCOR112N Expression |
| 16         | 68    | 13.7        | 4433   | EVCOR114N | EVCOR114N Expression |
| 17         | 68    | 13.7        | 7545   | AX840288  | AX840288 Sequence    |
| 18         | 68    | 13.7        | 11643  | AX840289  | AX840289 Sequence    |
| 19         | 67    | 13.5        | 273    | AY255709  | AY255709 Synthetic   |

|    |    |      |       |    |           |                      |
|----|----|------|-------|----|-----------|----------------------|
| 20 | 67 | 13.5 | 4032  | 6  | A71435    | A71435 Sequence 5    |
| 21 | 67 | 13.5 | 4032  | 6  | AR207453  | AR207453 Sequence    |
| 22 | 67 | 13.5 | 4032  | 6  | AR564463  | AR564463 Sequence    |
| 23 | 67 | 13.5 | 4032  | 6  | BD069509  | BD069509 Improved    |
| 24 | 67 | 13.5 | 4486  | 12 | EVCOR116N | EVCOR116N Expression |
| 25 | 63 | 12.7 | 399   | 6  | AR474264  | AR474264 Sequence    |
| 26 | 63 | 12.7 | 399   | 6  | AX430082  | AX430082 Sequence    |
| 27 | 62 | 12.4 | 6865  | 6  | AX093008  | AX093008 Sequence    |
| 28 | 62 | 12.4 | 10003 | 6  | AX093016  | AX093016 Sequence    |
| 29 | 62 | 12.4 | 10003 | 6  | AB021747  | AB021747 Oryza sat   |
| 30 | 62 | 12.4 | 7423  | 8  | AB021747  | AB021747 Oryza sat   |
| 31 | 62 | 12.4 | 7423  | 8  | AP003245  | AP003245 Oryza sat   |
| 32 | 62 | 12.4 | 7423  | 8  | AP003245  | AP003245 Oryza sat   |
| 33 | 62 | 12.4 | 7423  | 8  | AR474261  | AR474261 Sequence    |
| 34 | 62 | 12.4 | 7423  | 8  | AX430079  | AX430079 Sequence    |
| 35 | 62 | 12.4 | 7423  | 8  | AC091774  | AC091774 Oryza sat   |
| 36 | 62 | 12.4 | 7423  | 8  | AP003614  | AP003614 Oryza sat   |
| 37 | 62 | 12.4 | 7423  | 8  | AX342365  | AX342365 Sequence    |
| 38 | 62 | 12.4 | 7423  | 8  | AP005724  | AP005724 Oryza sat   |
| 39 | 62 | 12.4 | 7423  | 8  | AP006847  | AP006847 Oryza sat   |
| 40 | 62 | 12.4 | 7423  | 8  | AP003273  | AP003273 Oryza sat   |
| 41 | 62 | 12.4 | 7423  | 8  | AP005070  | AP005070 Oryza sat   |
| 42 | 62 | 12.4 | 7423  | 8  | AP006523  | AP006523 Oryza sat   |
| 43 | 62 | 12.4 | 7423  | 8  | AP003894  | AP003894 Oryza sat   |
| 44 | 62 | 12.4 | 7423  | 8  | AP005683  | AP005683 Oryza sat   |
| 45 | 62 | 12.4 | 7423  | 8  | AP006453  | AP006453 Oryza sat   |
|    | 62 | 12.4 | 7423  | 8  | AP003308  | AP003308 Oryza sat   |
|    | 62 | 12.4 | 7423  | 8  | AP003625  | AP003625 Oryza sat   |

ALIGNMENTS

|            |  |                                  |                 |                     |                 |
|------------|--|----------------------------------|-----------------|---------------------|-----------------|
| RESULT 1   | AX342368   | 498 bp                           | DNA             | linear              | PAT 12-JAN-2002 |
| LOCUS      | AX342368   | Sequence 7 from Patent EP167531. |                 |                     |                 |
| DEFINITION | AX342368   | AX342368.1 GI:18151811           |                 |                     |                 |
| ACCESSION  | AX342368   |                                  |                 |                     |                 |
| VERSION    | AX342368.1   | GI:18151811                      |                 |                     |                 |
| KEYWORDS   |  |                                  |                 |                     |                 |
| SOURCE     |  |                                  |                 |                     |                 |
| ORGANISM   |  |                                  |                 |                     |                 |
| REFERENCE  |  |                                  |                 |                     |                 |
| 1          | Behr, C.F., Hironaka, C., Heck, G.R. and You, J.                 |                                  |                 |                     |                 |
| ATTHORS    |  |                                  |                 |                     |                 |
| TITLE      |  |                                  |                 |                     |                 |
| JOURNAL    |  |                                  |                 |                     |                 |
| FEATURES   |  |                                  |                 |                     |                 |
| 1          | Behr, C.F., Hironaka, C., Heck, G.R. and You, J.                 |                                  |                 |                     |                 |
| 2          | Corn transformant pV-zmgf32 (hke03) and compositions and methods |                                  |                 |                     |                 |
| 3          | for detection thereof  |                                  |                 |                     |                 |
| 4          | Patent: EP 1167531-A 7 02-JAN-2002;                              |                                  |                 |                     |                 |
| 5          | Monsanto Technology LLC (US)                                     |                                  |                 |                     |                 |
| 6          | Location/Qualifiers  |                                  |                 |                     |                 |
| 7          | 1..498   |                                  |                 |                     |                 |
| 8          | /organism="synthetic construct"                                  |                                  |                 |                     |                 |
| 9          | /mol_type="unassigned DNA"                                       |                                  |                 |                     |                 |
| 10         | /db_xref="taxon:32630"   |                                  |                 |                     |                 |
| 11         | /note="1-304 Zea maize genomic DNA 305-349 construct             |                                  |                 |                     |                 |
| 12         | vector DNA 350-498 rice actin 1 promoter DNA"                    |                                  |                 |                     |                 |
| ORIGIN     |  |                                  |                 |                     |                 |
| 1          | ATTCAGTCCAAATCGGATTCAGGATTTACCGCCATGCTTTATGAGACAAGAGGCGAG        | 100.0%;                          | Score 498;      | DB 6;               | Length 498;     |
| 2          |  | Best local Similarity            | 100.0%;         | Pred. No. 8.2e-264; |                 |
| 3          |  | Matches 498;                     | Conservative 0; | Mismatches 0;       | Indels 0;       |
| 4          |  |                                  |                 |                     | Gaps 0;         |
| 5          |  |                                  |                 |                     |                 |
| 6          |  |                                  |                 |                     |                 |
| 7          |  |                                  |                 |                     |                 |
| 8          |  |                                  |                 |                     |                 |
| 9          |  |                                  |                 |                     |                 |
| 10         |  |                                  |                 |                     |                 |
| 11         |  |                                  |                 |                     |                 |
| 12         |  |                                  |                 |                     |                 |
| 13         |  |                                  |                 |                     |                 |
| 14         |  |                                  |                 |                     |                 |
| 15         |  |                                  |                 |                     |                 |
| 16         |  |                                  |                 |                     |                 |
| 17         |  |                                  |                 |                     |                 |
| 18         |  |                                  |                 |                     |                 |
| 19         |  |                                  |                 |                     |                 |
| 20         |  |                                  |                 |                     |                 |
| 21         |  |                                  |                 |                     |                 |
| 22         |  |                                  |                 |                     |                 |
| 23         |  |                                  |                 |                     |                 |
| 24         |  |                                  |                 |                     |                 |
| 25         |  |                                  |                 |                     |                 |
| 26         |  |                                  |                 |                     |                 |
| 27         |  |                                  |                 |                     |                 |
| 28         |  |                                  |                 |                     |                 |
| 29         |  |                                  |                 |                     |                 |
| 30         |  |                                  |                 |                     |                 |
| 31         |  |                                  |                 |                     |                 |
| 32         |  |                                  |                 |                     |                 |
| 33         |  |                                  |                 |                     |                 |
| 34         |  |                                  |                 |                     |                 |
| 35         |  |                                  |                 |                     |                 |
| 36         |  |                                  |                 |                     |                 |
| 37         |  |                                  |                 |                     |                 |
| 38         |  |                                  |                 |                     |                 |
| 39         |  |                                  |                 |                     |                 |
| 40         |  |                                  |                 |                     |                 |
| 41         |  |                                  |                 |                     |                 |
| 42         |  |                                  |                 |                     |                 |
| 43         |  |                                  |                 |                     |                 |
| 44         |  |                                  |                 |                     |                 |
| 45         |  |                                  |                 |                     |                 |



Db 181 AGCCGCGGTGCGCATTCGCGCAGACTCCTCTCTCGGCATGAGCCGATCTTTCTCTG 240  
Qy 241 GCATTTCCAAACCTTAGAGAGTGGTCCCTGGTGGGCTCGGCGCAGCAAGCTTAG 300  
Db 241 GCATTTCCAAACCTTAGAGAGTGGTCCCTGGTGGGCTCGGCGCAGCAAGCTTAG 300  
Qy 301 CGGCCCCGCGGTGATCCAGCTTATCCCTAGGCGCGCGCTTACCAAGCTTACTC 360  
Db 301 CGGCCCCGCGGTGATCCAGCTTATCCCTAGGCGCGCGCTTACCAAGCTTACTC 360  
Qy 361 GAGGTCAATCATATGCTTGAAGAGAGTGGATAGTCCAAATTAACCAAGGTAGA 420  
Db 361 GAGGTCAATCATATGCTTGAAGAGAGTGGATAGTCCAAATTAACCAAGGTAGA 420  
Qy 421 TTACCGGTCAAAAGTGAACATCATGTTAAAGGTATTAAGTAAATATCGTAAATA 480  
Db 421 TTACCGGTCAAAAGTGAACATCATGTTAAAGGTATTAAGTAAATATCGTAAATA 480  
Qy 481 AAGGTGCGCCCAAGTGAA 498  
Db 481 AAGGTGCGCCCAAGTGAA 498

RESULT 2  
AX342369 1183 bp DNA linear PAT 12-JAN-2002

LOCUS AX342369 Sequence 8 from Patent EP1167531.

DEFINITION AX342369

ACCESSION AX342369.1 GI:18151812

VERSION AX342369.1 GI:18151812

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS Behr, C.F., Hironaka, C., Heck, G.R. and You, J.

TITLE Corn transformant pv-zmgt32 (nke603) and compositions and methods

JOURNAL for detection thereof

Patent: EP 1167531-A 8 02-JAN-2002;

Monasanto Technology LLC (US)

FEATURES

source

1. .1183

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="1-164 Agrobacterium tumefaciens nos 3' terminator

165-381 construct vector DNA 382-686 Zea mays plasmid

genes, tpst1 and tpst2 687-1183 Zea mays genomic DNA"

ORIGIN

Query Match 23.7%; Score 118; DB 6; Length 1183;

Best Local Similarity 100.0%; Pred. No. 1.8e-53;

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 308 CGCGTGTACCAAGCTTATCCCTAGGCGCGCGCTTACCAAGCTTATCGAGTCA 367

Db 381 CGCGTGTACCAAGCTTATCCCTAGGCGCGCGCTTACCAAGCTTATCGAGTCA 322

Qy 368 TTGATATGCTTGAAGAGAGTGGATAGTCCAAATTAACCAAGGTAGAATACC 425

Db 321 TTGATATGCTTGAAGAGAGTGGATAGTCCAAATTAACCAAGGTAGAATACC 264

RESULT 3

AR229555 2378 bp DNA linear PAT 20-DEC-2002

LOCUS AR229555 Sequence 27 from Patent US 6448476.

DEFINITION AR229555

ACCESSION AR229555

VERSION AR229555.1 GI:27269171

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2378)

AUTHORS Barry, G.F.

TITLE Plants and plant cells transformation to express an

JOURNAL AMPA-N-acetyltransferase

Patent: US 6448476-A 27 10-SEP-2002;

Location/Qualifiers

1. .2378

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Query Match 20.3%; Score 101; DB 6; Length 2378;

Best Local Similarity 100.0%; Pred. No. 4.7e-44;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 GATATCCCTAGGCGCGCGCTTACCAAGCTTATCGAGTCAATATGCTTGAAG 384

Db 1 GATATCCCTAGGCGCGCGCTTACCAAGCTTATCGAGTCAATATGCTTGAAG 60

Qy 385 AGAGTCGGATATGTCCTCAAAATTAACCAAGGTAGAATACC 425

Db 61 AGAGTCGGATATGTCCTCAAAATTAACCAAGGTAGAATACC 101

RESULT 4

S44221 1266 bp DNA linear PLN 18-MAR-1998

LOCUS S44221 Act1-actin 1 (5' region, promoter) [rice, Genomic, 1266 nt].

DEFINITION S44221

ACCESSION S44221.1 GI:256131

VERSION S44221.1 GI:256131

KEYWORDS

SOURCE

ORGANISM

Oryza sativa

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS McElroy, D., Zhang, W., Cao, J. and Wu, R.

TITLE Isolation of an efficient actin promoter for use in rice

transformation

JOURNAL Plant Cell 2 (2), 163-171 (1990)

MEDLINE 92404732

PUBMED 2136653

REMARK GenBank staff at the National Library of Medicine created this

entry [NCBI gibbs 114121] from the original journal article.

This sequence comes from Fig. 2.

FEATURES

source

1. .1266

/organism="Oryza sativa"

/mol\_type="genomic DNA"

/db\_xref="taxon:4530"

1246. .1266

/gene="Act1"

/note="actin 1"

1246. .1266

/gene="Act1"

/note="actin 1"

/codon\_start=1

/protein\_id="AA013837.1"

/db\_xref="GI:4261543"

/translation="MADAEDI"

ORIGIN

Query Match 14.1%; Score 70; DB 8; Length 1266;

Best Local Similarity 100.0%; Pred. No. 6.7e-27;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 356 TACTCGAGTCATTCATATGCTTGAAGAGAGTCCGAGATGTCCAAAATTAACCAAG 415

Db 10 TACTCGAGTCATTCATATGCTTGAAGAGAGTCCGAGATGTCCAAAATTAACCAAG 69

Qy 416 TAAGATTACC 425



Db 70 TAAGATTACC 79

RESULT 5  
LOCUS 150115 1404 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 7 from patent US 5641876.  
ACCESSION 150115  
VERSION 150115.1 GI:2472335  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1404)  
AUTHORS McElroy,D. and Wu,R.  
TITLE Rice actin gene and promoter  
JOURNAL Patent: US 5641876-A 7 24-JUN-1997;  
FEATURES  
source  
1. .1404  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 14.1%; Score 70; DB 6; Length 1404;  
Best Local Similarity 100.0%; Pred. No. 6.7e-27;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 415  
|||||  
9 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 68  
|||||

Db 416 TAAGATTACC 425  
|||||  
69 TAAGATTACC 78  
|||||

RESULT 6  
LOCUS AX044094 1501 bp DNA linear PAT 24-NOV-2000  
DEFINITION Sequence 49 from Patent W00066748.  
ACCESSION AX044094  
VERSION AX044094.1 GI:11342972  
KEYWORDS  
SOURCE Oryza sp.  
ORGANISM Oryza sp.  
REFERENCE 1  
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and  
TITLE Herbicide resistant plants  
JOURNAL Patent: WO 0066748-A 49 09-NOV-2000;  
FEATURES  
source  
1. .1501  
/organism="Oryza sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:52841"

ORIGIN

Query Match 14.1%; Score 70; DB 6; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 6.7e-27;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 415  
|||||  
750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 809  
|||||

QY 416 TAAGATTACC 425  
|||||  
810 TAAGATTACC 819  
|||||

RESULT 7  
LOCUS AX044144 1501 bp DNA linear PAT 24-NOV-2000  
DEFINITION Sequence 44 from Patent W00066747.  
ACCESSION AX044144  
VERSION AX044144.1 GI:11343022  
KEYWORDS  
SOURCE Oryza sp.  
ORGANISM Oryza sp.  
REFERENCE 1  
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and  
TITLE Herbicide resistant plants  
JOURNAL Patent: WO 0066747-A 44 09-NOV-2000;  
FEATURES  
source  
1. .1501  
/organism="Oryza sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:52841"

ORIGIN

Query Match 14.1%; Score 70; DB 6; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 6.7e-27;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 415  
|||||  
750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 809  
|||||

Db 416 TAAGATTACC 425  
|||||  
810 TAAGATTACC 819  
|||||

RESULT 8  
LOCUS 150113 2199 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 5 from patent US 5641876.  
ACCESSION 150113  
VERSION 150113.1 GI:2472333  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2199)  
AUTHORS McElroy,D. and Wu,R.  
TITLE Rice actin gene and promoter  
JOURNAL Patent: US 5641876-A 5 24-JUN-1997;  
FEATURES  
source  
1. .2199  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 14.1%; Score 70; DB 6; Length 2199;  
Best Local Similarity 100.0%; Pred. No. 6.8e-27;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 415  
|||||  
806 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAACAAGG 865  
|||||

QY 416 TAAGATTACC 425  
|||||  
866 TAAGATTACC 875  
|||||

RESULT 9

|   |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|---|---|---|-------------|-----|--------------------------|---------------------|--|---------|-------------------------------------|-----------------------|--|-------------------------|--|-----------|---|-----------|--|----------------|--|---------------------------|--|--------------------------|--|------------------------|--|---|--|--|--|--|--|---|--|---|--|------------|--|------------|------------------------------------|------------|------------------------------------|------------|----------------|--|--|------------|---------------|------------|---------------|--------------|--|----------------------------|--|----------------|--|---------------------|--|--------------------------|--|------------------------|--|---|--|--|--|--|--|---|--|-----------------------|--|------------|------------------------|------|--|
|   | LOCUS   | Ay452735  | 4895 bp     | DNA | circular SYN 19-DEC-2003 |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | DEFINITION  | Reporter vector pActcXN, complete sequence.   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | ACCESSION   | AY452735  |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | VERSION   | AY452735.1  | GI:39636979 |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | KEYWORDS  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | SOURCE ORGANISM   | ' Reporter vector pActcXN<br>Reporter vector pActcXN<br>Other sequences: artificial sequences; vectors.'  |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | REFERENCE AUTHORS   | Vickers,C.E., Xue,G.P. and Greshoff,P.M.<br>A synthetic xylanase as a novel reporter in plants  |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | TITLE   | Plant Cell Rep. 22 (2), 135-140 (2003)  |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | JOURNAL MEDLINE   | 2267549   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | PUBMED  | 12845475  |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | REFERENCE   | 2 (bases 1 to 4895)<br><br>Vickers,C.E.<br>Direct Submission<br>Submitted (29-Oct-2003) ARC Centre for Integrative Legume Research,<br>The University of Queensland, Room 213, John Hines Building (69) ,<br>St. Lucia, QLD 4072, Australia   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | FEATURES  | <table border="0"> <tr><td>Location/Qualifiers</td><td></td></tr> <tr><td>    1..4895</td><td>/organism="Reporter vector pActcXN"</td></tr> <tr><td>        /mol_type="other DNA"</td><td></td></tr> <tr><td>         /db_xref="taxon:255329"</td><td></td></tr> <tr><td>       267..2274</td><td>/note="sequence from cloning vector pBR322"</td></tr> <tr><td>       468..1328</td><td>/note="ampicillin resistance protein; bla"</td></tr> <tr><td>         /codon_start=1</td><td></td></tr> <tr><td>         /product="beta-lactamase"</td><td></td></tr> <tr><td>         /protein_id="AAK29085.1"</td><td></td></tr> <tr><td>         /db_xref="gi:39636980"</td><td></td></tr> <tr><td>         /translat="MSIOHPVALIPFPAFCIPIVAHPEITLVKYDAEDQGVGVC</td><td></td></tr> <tr><td>         IELDNSKKILSEFRPERPMMSTPKULIGAVYSIDAOEOLGRHYSONDIYE</td><td></td></tr> <tr><td>         YSPVEKHLDGMTREICSAFITMSDNTANLLTTGGKEFLAHNGDVTKRL</td><td></td></tr> <tr><td>         DRWEELNEAIENDRDTTPVAMATTLRKLTGTGLTIVASROOIMWENADKVAGPIP</td><td></td></tr> <tr><td>         LRSALPGMFIADKSAGEBSRGIIIALGPDKPSRIIVIYTGSQAINDERRQLIA</td><td></td></tr> <tr><td>         EGASLIKHW"</td><td></td></tr> <tr><td>       2345..2367</td><td>/note="r7 RNA polymerase promoter"</td></tr> <tr><td>       2426..3338</td><td>/note="from rice actin gene; Act1"</td></tr> <tr><td>       3339..3792</td><td>/note="sgvnt2"</td></tr> <tr><td>       3800.."inttron 1 from rice actin gene; Act1"</td><td></td></tr> <tr><td>       3800..4513</td><td>/gene="sxynA"</td></tr> <tr><td>       3800..4513</td><td>/gene="sxynA"</td></tr> <tr><td>       /gene="synA"</td><td></td></tr> <tr><td>       /note="synthetic xylanase"</td><td></td></tr> <tr><td>       /codon_start=1</td><td></td></tr> <tr><td>       /product="Xylanase"</td><td></td></tr> <tr><td>       /protein_id="AAK29086.1"</td><td></td></tr> <tr><td>       /db_xref="gi:39636981"</td><td></td></tr> <tr><td>       /translation="MASNGKFVTGNQGONHQHKVDGFSYSTIWLNDTGSGSMVLGSG</td><td></td></tr> <tr><td>       ATEFKEMMAAVVRGNFLARGLDFGSSOKAKDYDDYLGADVATIKQTASAGSNRLCV</td><td></td></tr> <tr><td>       YGWFNRGIRJDEPLIVEYYIIEBWDMVMWDAPCGKWTVIDGAQYKIIFOMDHPTPINGCS</td><td></td></tr> <tr><td>       ETPFKOYFSVKOOKRTSHGITWSDFHEWKAWKGWGIGNLYEVALLNAEWQSOGSVADVTLL</td><td></td></tr> <tr><td>       LDVVTFPPKSGSPATSAAAR"</td><td></td></tr> <tr><td>       4535..4872</td><td>/note="from rbcS gene"</td></tr> <tr><td>       4895</td><td>/note="sp6 RNA polymerase transcription initiation site"</td></tr> </table> |             |     |                          | Location/Qualifiers |  | 1..4895 | /organism="Reporter vector pActcXN" | /mol_type="other DNA" |  | /db_xref="taxon:255329" |  | 267..2274 | /note="sequence from cloning vector pBR322" | 468..1328 | /note="ampicillin resistance protein; bla" | /codon_start=1 |  | /product="beta-lactamase" |  | /protein_id="AAK29085.1" |  | /db_xref="gi:39636980" |  | /translat="MSIOHPVALIPFPAFCIPIVAHPEITLVKYDAEDQGVGVC |  | IELDNSKKILSEFRPERPMMSTPKULIGAVYSIDAOEOLGRHYSONDIYE |  | YSPVEKHLDGMTREICSAFITMSDNTANLLTTGGKEFLAHNGDVTKRL |  | DRWEELNEAIENDRDTTPVAMATTLRKLTGTGLTIVASROOIMWENADKVAGPIP |  | LRSALPGMFIADKSAGEBSRGIIIALGPDKPSRIIVIYTGSQAINDERRQLIA |  | EGASLIKHW" |  | 2345..2367 | /note="r7 RNA polymerase promoter" | 2426..3338 | /note="from rice actin gene; Act1" | 3339..3792 | /note="sgvnt2" | 3800.."inttron 1 from rice actin gene; Act1" |  | 3800..4513 | /gene="sxynA" | 3800..4513 | /gene="sxynA" | /gene="synA" |  | /note="synthetic xylanase" |  | /codon_start=1 |  | /product="Xylanase" |  | /protein_id="AAK29086.1" |  | /db_xref="gi:39636981" |  | /translation="MASNGKFVTGNQGONHQHKVDGFSYSTIWLNDTGSGSMVLGSG |  | ATEFKEMMAAVVRGNFLARGLDFGSSOKAKDYDDYLGADVATIKQTASAGSNRLCV |  | YGWFNRGIRJDEPLIVEYYIIEBWDMVMWDAPCGKWTVIDGAQYKIIFOMDHPTPINGCS |  | ETPFKOYFSVKOOKRTSHGITWSDFHEWKAWKGWGIGNLYEVALLNAEWQSOGSVADVTLL |  | LDVVTFPPKSGSPATSAAAR" |  | 4535..4872 | /note="from rbcS gene" | 4895 | /note="sp6 RNA polymerase transcription initiation site" |
| Location/Qualifiers   |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 1..4895   | /organism="Reporter vector pActcXN"                               |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /mol_type="other DNA"   |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /db_xref="taxon:255329"                                       |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 267..2274   | /note="sequence from cloning vector pBR322"                       |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 468..1328   | /note="ampicillin resistance protein; bla"                        |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /codon_start=1  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /product="beta-lactamase"                                     |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /protein_id="AAK29085.1"                                      |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /db_xref="gi:39636980"  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /translat="MSIOHPVALIPFPAFCIPIVAHPEITLVKYDAEDQGVGVC           |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| IELDNSKKILSEFRPERPMMSTPKULIGAVYSIDAOEOLGRHYSONDIYE            |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| YSPVEKHLDGMTREICSAFITMSDNTANLLTTGGKEFLAHNGDVTKRL              |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| DRWEELNEAIENDRDTTPVAMATTLRKLTGTGLTIVASROOIMWENADKVAGPIP       |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| LRSALPGMFIADKSAGEBSRGIIIALGPDKPSRIIVIYTGSQAINDERRQLIA         |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| EGASLIKHW"  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 2345..2367  | /note="r7 RNA polymerase promoter"                                |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 2426..3338  | /note="from rice actin gene; Act1"                                |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 3339..3792  | /note="sgvnt2"  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 3800.."inttron 1 from rice actin gene; Act1"                  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 3800..4513  | /gene="sxynA"   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 3800..4513  | /gene="sxynA"   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /gene="synA"  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /note="synthetic xylanase"                                    |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /codon_start=1  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /product="Xylanase"   |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /protein_id="AAK29086.1"                                      |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /db_xref="gi:39636981"  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| /translation="MASNGKFVTGNQGONHQHKVDGFSYSTIWLNDTGSGSMVLGSG     |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| ATEFKEMMAAVVRGNFLARGLDFGSSOKAKDYDDYLGADVATIKQTASAGSNRLCV      |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| YGWFNRGIRJDEPLIVEYYIIEBWDMVMWDAPCGKWTVIDGAQYKIIFOMDHPTPINGCS  |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| ETPFKOYFSVKOOKRTSHGITWSDFHEWKAWKGWGIGNLYEVALLNAEWQSOGSVADVTLL |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| LDVVTFPPKSGSPATSAAAR"   |   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 4535..4872  | /note="from rbcS gene"  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| 4895  | /note="sp6 RNA polymerase transcription initiation site"          |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | CDS   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | misc_feature  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | promoter  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | intron  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | gene  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | CDS   |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | misc_feature  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | terminator  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | misc_feature  |   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
|   | ORIGIN  | Query Match      14.1%; Score 70; DB 12; Length 4895;<br>Best Local Similarity    100.0%; Pred.No. 6.9e-27;<br>Matches      70; Conservative    0; Mismatched -0; Indels    0; Gaps    0;   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| OY  | 356 TACTCAGAGTCATTATAATGCTTAGAGAAGAGTGC GGATAGTCCAAAATTAACAACAGG  | 415   |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |
| DB  | 2423 TAATCAAGCATTTATATATGCTTTGAGAGAGAGTGGGATATCTCCAATATTAACAACAGG | 2482  |             |     |                          |                     |  |         |                                     |                       |  |                         |  |           |   |           |  |                |  |                           |  |                          |  |                        |  |   |  |  |  |  |  |   |  |   |  |            |  |            |                                    |            |                                    |            |                |  |  |            |               |            |               |              |  |                            |  |                |  |                     |  |                          |  |                        |  |   |  |  |  |  |  |   |  |                       |  |            |                        |      |  |

| RESULT 10             | 150112   | 5643 bp     | DNA | linear | PAT 07-OCT-1997 |
|-----------------------|--|-------------|-----|--------|-----------------|
| LOCUS                 | Sequence 4 from patent US 5641876.                                   |             |     |        |                 |
| DEFINITION            | 150112   |             |     |        |                 |
| ACCESSION             | 150112   |             |     |        |                 |
| VERSION               | 150112.1   | GI:2472332  |     |        |                 |
| KEYWORDS              |  |             |     |        |                 |
| SOURCE                | Unknown.   |             |     |        |                 |
| ORGANISM              | Unknown.   |             |     |        |                 |
| REFERENCE             | Unclassified.  |             |     |        |                 |
| AUTHORS               | 1 (bases 1 to 5643)  |             |     |        |                 |
| TITLE                 | McElroy, D. and Wu, R.   |             |     |        |                 |
| JOURNAL               | Rice <i>activator gene</i> and promoter                              |             |     |        |                 |
| FEATURES              | Patent: US 5641876-A 4 24-JUN-1997;                                  |             |     |        |                 |
| SOURCE                | Location: <i>Genetic Engineering</i>                                 |             |     |        |                 |
| ORIGIN                | 1..5643  |             |     |        |                 |
|                       | /organism="unknown"  |             |     |        |                 |
|                       | /mol_type="unassigned DNA"   |             |     |        |                 |
| Query Match           | 14.1%; Score 70; DB 6; Length 5643;                                  |             |     |        |                 |
| Best Local Similarity | 100.0%; Pred. No. 6.9e-27;   |             |     |        |                 |
| Matches               | 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;                  |             |     |        |                 |
| OY                    | 356 TACTCGAGGTCATTGCTTGAAGAGAGAGTCGGGATGTCCTCAAAATTAACAAAGG 415      |             |     |        |                 |
| DB                    | 809 TACTCGAGGTCATTGCTTGAAGAGAGAGTCGGGATGTCCTCAAAATTAACAAAGG 868      |             |     |        |                 |
| OY                    | 416 TTAAGATTACC 425  |             |     |        |                 |
| DB                    | 869 TTAAGATTACC 878  |             |     |        |                 |
| RESULT 11             |  |             |     |        |                 |
| LOCUS                 | AX384394   | 9359 bp     | DNA | linear | PAT 19-MAR-2002 |
| DEFINITION            | Sequence 3 from Patent WO0214524.                                    |             |     |        |                 |
| ACCESSION             | AX384394   |             |     |        |                 |
| VERSION               | AX384394.1   | GI:19577678 |     |        |                 |
| KEYWORDS              |  |             |     |        |                 |
| SOURCE                |  |             |     |        |                 |
| ORGANISM              | Synthetic construct  |             |     |        |                 |
| REFERENCE             | Synthetic construct  |             |     |        |                 |
| AUTHORS               | other sequences, artificial sequences.                               |             |     |        |                 |
| TITLE                 | 1  |             |     |        |                 |
| JOURNAL               | van der Valk, P., van Dun, C.M., Smeekens, S.C. and Proveniers, M.C. |             |     |        |                 |
| FEATURES              | Inhibition of generative propagation in genetically modified         |             |     |        |                 |
| SOURCE                | herbicide resistant grasses  |             |     |        |                 |
|                       | Patent: WO 0214524-A 3 21-FEB-2002;                                  |             |     |        |                 |
|                       | Advanta Seeds B.V. (NL)  |             |     |        |                 |
|                       | Location/Qualifiers  |             |     |        |                 |
|                       | 1..9359  |             |     |        |                 |
|                       | /organism="synthetic construct"                                      |             |     |        |                 |
|                       | /mol_type="unassigned DNA"   |             |     |        |                 |
|                       | /db_xref="taxon:32630"   |             |     |        |                 |
|                       | 839..1699  |             |     |        |                 |
|                       | /note="unamed protein product; Beta-lactamase gene                   |             |     |        |                 |
|                       | (Ampr)"  |             |     |        |                 |
|                       | /codon_start=1   |             |     |        |                 |
|                       | /transl_table=1  |             |     |        |                 |
|                       | /protein_id="CAD28571.1"   |             |     |        |                 |
|                       | /db_xref="GI:19577679"   |             |     |        |                 |
|                       | /translation="MSIOHERVALIPPEAFCLEVPAPPETLVKDAEDQLGARVGY              |             |     |        |                 |
|                       | IELDNGKLTIESFRPEREPFPMSTPKVLGAVLSRIDGOEOLGRIRHSNDLVE                 |             |     |        |                 |
|                       | YSPTKRGKLDGKTVRELCSAATMSNTNANLLTIGGPKELTALAHNGDHVTE                  |             |     |        |                 |
|                       | DRMPELNEALPNDERTTTPVMAATTIRKLLTSELTLASROQLIMMEADKTAGEL               |             |     |        |                 |
|                       | LRSLALPGWFLADKSGAGERSRGIIALGDPKPSRI VLYITTSQATMDERNRQIA              |             |     |        |                 |
|                       | ETGSLIKHW"   |             |     |        |                 |

promoter 2941..4920  
 /note="Ubi-promoter from maize"  
 misc\_feature 4921..6400  
 /note="ACH1 gene from Arabidopsis thaliana"  
 polyA\_signal 6401..6672  
 /note="Poly-A signal from the nopaline synthetase gene from Agrobacterium tumefaciens"  
 misc\_feature 7434..8084  
 /note="First exon-intron combination from Ubi-maize"  
 misc\_feature 8085..9119  
 /note="Hygromycin resistance gene from Escherichia coli"  
 polyA\_signal 9120..9359

Query Match 14.1%; Score 70; DB 6; Length 9359;  
 Best Local Similarity 100.0%; Pred. No. 7e-27;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATTCCTTGAGAGAGAGTCGGGATGATCCAAATATAACCAAGG 415  
 |||||  
 DB 6688 TACTCGAGTCATTCATTCCTTGAGAGAGAGTCGGGATGATCCAAATATAACCAAGG 6747

QY 416 TAAAGTTACC 425  
 |||||

DB 6748 TAAAGTTACC 6757

RESULT 12  
 AX473364 9359 bp DNA linear PAT 09-AUG-2002  
 LOCUS AX473364  
 DEFINITION Sequence 1 from Patent W00214486.  
 ACCESSION AX473364  
 VERSION AX473364.1 GI:22207993  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.

REFERENCE 1  
 AUTHORS van der Valk, P., van Dun, C.M., Smeekens, S.C. and Proveniers, M.C.  
 TITLE Inhibition of generative propagation in genetically modified grasses  
 JOURNAL Patent: WO 0214486-A 1 21-FEB-2002;  
 Advanta Seeds B.V. (NL)  
 FEATURES  
 source Location/Qualifiers  
 1..9359  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 promoter 2941..4920  
 /note="Ubi-promoter from maize"  
 polyA\_signal 6401..6672  
 /note="Poly-A signal from the nopaline synthetase gene from Agrobacterium tumefaciens"  
 misc\_feature 7434..8084  
 /note="First exon-intron combination from Ubi-maize"  
 polyA\_signal 9120..9359

ORIGIN  
 Query Match 14.1%; Score 70; DB 6; Length 9359;  
 Best Local Similarity 100.0%; Pred. No. 7e-27;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATTCCTTGAGAGAGAGTCGGGATGATCCAAATATAACCAAGG 415  
 |||||  
 DB 6688 TACTCGAGTCATTCATTCCTTGAGAGAGAGTCGGGATGATCCAAATATAACCAAGG 6747

QY 416 TAAAGTTACC 425  
 |||||

DB 6748 TAAAGTTACC 6757

RESULT 13  
 AC091532

LOCUS AC091532 147344 bp DNA linear PLN 11-FEB-2003  
 DEFINITION Oryza sativa chromosome 3 BAC OSJNBa0078A17 genomic sequence,  
 complete sequence.  
 AC091532  
 VERSION AC091532.13 GI:28173090  
 KEYWORDS HTG.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 147344)  
 Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, R.M.,  
 Overton, II, L.L., Tsirint, T., Kim, M.M., Bera, J.J., Jin, S.S.,  
 Padrosch, D.W., Tallon, L.J., Koo, H., Ziemann, V., Heiao, J., Blunt, S.,  
 Vanaken, S.S., Riedmuller, S.B., Utecherback, T.T., Feldlyum, T.V.,  
 Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.,  
 White, O., Salzberg, S.L. and Fraser, C.M.  
 Oryza sativa chromosome 3 BAC OSJNBa0078A17 genomic sequence  
 Unpublished  
 2 (bases 1 to 147344)  
 Buell, R.  
 Direct Submission  
 Submitted (03-MAY-2001) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 3 (bases 1 to 147344)  
 Buell, R.  
 Direct Submission  
 Submitted (12-NOV-2002) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 4 (bases 1 to 147344)  
 Buell, R.  
 Direct Submission  
 Submitted (31-JAN-2003) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 5 (bases 1 to 147344)  
 Buell, R.  
 Direct Submission  
 Submitted (11-FEB-2003) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA, rhuels@igrr.org  
 On Jan 31, 2003 this sequence version replaced gi:24308660.  
 Address all correspondence to: rice@igrr.org

COMMENT  
 BAC clone OSJNBa0078A17 is from Oryza sativa chromosome 3  
 The orientation of the sequence is from SP6 to T7 end of the BAC  
 clone.  
 Genes were identified by a combination of several methods: Gene  
 prediction programs including Fgenesh (<http://www.softberry.com/>),  
 genscan and GenScan+ (Chris Burge,  
<http://CCR-081.mit.edu/GENSCAN.html>), GenMarkES (Mark Borodovsky,  
<http://genemark.biology.gatech.edu/genemark/>), and GeneSplicer  
 (Mihaila Petrea and Steven Salzberg, [contact mperleatig.org](http://contact.mperleatig.org)),  
 searches of the complete sequence against a peptide database and  
 the plant EST database at TIGR (<http://www.tigr.org/tdb/tgl.shtml>).  
 Annotated genes are named to indicate the level of evidence for  
 their annotation. Genes with similarity to other proteins are named  
 after the database hits. Genes without significant peptide  
 similarity but with EST similarity are named as unknown proteins.  
 Genes without protein or EST similarity, that are predicted by more  
 than two gene prediction programs over most of their length are  
 annotated as hypothetical proteins. Genes encoding tRNAs are  
 predicted by tRNAscan-SE (Sean Eddy,  
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are  
 identified by repeatmasker (Arian Smit,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES  
 source  
 1..147344  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"

This BAC overlaps with rice BACOSJNBa0014G15 (AC090882).



```

VERSION      150114.1  GI:2472334
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    Unclassified.
AUTHORS      1 (bases 1 to 1392)
TITLE        McElroy, D. and Wu, R.
JOURNAL      Rice actin gene and promoter
              Patent: US 5641876-A 6 24-JUN-1997;
              Location/Qualifiers
FEATURES
  source      1..1392
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN
Query Match      13.7%; Score 68; DB 6; Length 1392;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358  CTGAGGTCATTGCTTGAAGAGAGTCGGATGCTCCAAATAAACAAGGTA 417
          |||||||
          1  CTGAGGTCATTGCTTGAAGAGAGTCGGATGCTCCAAATAAACAAGGTA 60

DB      418  AGATTACC 425
          |||||||
          61  AGATTACC 68

RESULT 15
LOCUS       EVCOR112N      4412 bp      DNA      linear      SYN 18-MAR-1996
DEFINITION  Expression vector pCOR112N (modified from pCOR112 in [2]).
ACCESSION   Z29587
VERSION     Z29587.1  GI:452344
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS      Depicker, A., Stachel, S., Dhaese, P., Zambrzyki, P. and Goodman, H.M.
TITLE        Nopaline synthase: transcript mapping and DNA sequence
JOURNAL      J. Mol. Appl. Genet. 1 (6), 561-573 (1982)
MEDLINE     83110651
PUBMED      7153689

REFERENCE
AUTHORS      2
              McElroy, D., Blowers, A.D., Jones, B. and Wu, R.
TITLE        Construction of expression vectors based on the rice actin 1 (Act1)
              5' region for use in monocot transformation
JOURNAL      Mol. Gen. Genet. 231 (1), 150-160 (1991)
MEDLINE     92092956
PUBMED      1753941
REMARK      (sites)
REFERENCE    3 (bases 1 to 4412)
AUTHORS      Liu, L., Dasgupta, L., Davies, J.W. and Hull, R.
TITLE        Modified vectors for monocot transformation toward virus resistance
JOURNAL      Unpublished
REFERENCE    4 (bases 1 to 4412)
AUTHORS      Liu, L.
TITLE        Direct Submision
              Submitted (26-JAN-1994) Liu L., John Innes Institute, Virus
              Research, Colney Lane, Norwich, United Kingdom, NR4 7UH
              Location/Qualifiers
FEATURES
  source      1..4412
              /organism="synthetic construct"
              /mol_type="other DNA"
              /db_xref="taxon:32630"
              1..30
              /note="multiple cloning sites"
              3..5
              /note="start codon for gene expression"
              27..282
              /note="nos terminator"

```

```

misc_feature    /citation=[1]
                 /note=".962
                 /note="phage fl region"
CDS             1093..1953
                 /codon_start=1
                 /transl_table=11
                 /product="beta-lactamase"
                 /protein_id="CAA82704.1"
                 /db_xref="GI:452345"
                 /translation="MSIQHFRVALIPFPAAFCGLPVPAPHPETLVKVKDAEDQLGARVGY
                 IELDNLNGKILSEFRPEERFPMWSTFKVLGCAVLSRIDAGQQLGRIRIHYSQNDLVE
                 VSPVTEKHLTDGMTVRELCSAAITWSDNTAANLLITIGQPKELTAFLNMGDVTRL
                 DRWPELNAELPNDERDTTPVMAATTLAKLLTGSLTLASQQQLIDMEADKVAQPL
                 LRSALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIIVIIYTTGSQATWDERNRQIA
                 EIGASLIKHM"
                 1954..2894
                 /note="COLE1 origin"
                 3181..4412
                 /note="rice actin 1 5' region (act1 promoter)"
                 /citation=[2]

ORIGIN
Query Match      13.7%; Score 68; DB 12; Length 4412;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358  CTGAGGTCATTGCTTGAAGAGAGTCGGATGCTCCAAATAAACAAGGTA 417
          |||||||
          DB      3181  CTGAGGTCATTGCTTGAAGAGAGTCGGATGCTCCAAATAAACAAGGTA 3240

QY      418  AGATTACC 425
          |||||||
          DB      3241  AGATTACC 3248

Search completed: February 10, 2005, 13:13:39
Job time : 3075.76 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 111.027 Seconds  
(without alignments)  
7855.666 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18  
Sequence: 1 TGCCTGCTGCTGCTGCTT 18

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:\*  
1: gb\_Da:\*  
2: gb\_Htg:\*  
3: gb\_In:\*  
4: gb\_Om:\*  
5: gb\_Ov:\*  
6: gb\_Pat:\*  
7: gb\_Pn:\*  
8: gb\_Pl:\*  
9: gb\_Pr:\*  
10: gb\_Ro:\*  
11: gb\_Sls:\*  
12: gb\_Sy:\*  
13: gb\_Un:\*  
14: gb\_Vi:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 18    | 100.0       | 18     | 6     | AX342372 Sequence  |
| 2          | 18    | 100.0       | 1183   | 6     | AX342369 Sequence  |
| 3          | 18    | 100.0       | 2386   | 10    | BC055684 Mus muscu |
| 4          | 18    | 100.0       | 2699   | 10    | BC034572 Mus muscu |
| 5          | 18    | 100.0       | 2835   | 10    | BC027167 Mus muscu |
| 6          | 18    | 100.0       | 3788   | 10    | AY341877 Mus muscu |
| 7          | 18    | 100.0       | 175440 | 10    | AC125374 Mus muscu |
| 8          | 18    | 100.0       | 218326 | 2     | AC094355 Rattus no |
| 9          | 17    | 94.4        | 140256 | 10    | AC132091 Mus muscu |
| 10         | 17    | 94.4        | 164405 | 9     | AC016256 Homo sapi |
| 11         | 17    | 94.4        | 172148 | 3     | U00000.1 Leishmani |
| 12         | 17    | 94.4        | 175062 | 10    | AC138156 Mus muscu |
| 13         | 17    | 94.4        | 229391 | 2     | AC105823 Rattus no |
| 14         | 17    | 94.4        | 231074 | 2     | AC142509 Rattus no |
| 15         | 17    | 94.4        | 231140 | 2     | AC097799 Rattus no |
| 16         | 17    | 94.4        | 233915 | 10    | AC135290 Mus muscu |
| 17         | 17    | 94.4        | 251289 | 2     | AC095160 Rattus no |
| 18         | 17    | 94.4        | 256017 | 2     | AC121398 Rattus no |
| 19         | 16    | 88.9        | 359    | 8     | AY301027 Zea mays  |

|    |    |      |     |   |                   |
|----|----|------|-----|---|-------------------|
| 20 | 16 | 88.9 | 359 | 8 | AY301036 Zea mays |
| 21 | 16 | 88.9 | 365 | 8 | AY301028 Zea mays |
| 22 | 16 | 88.9 | 365 | 8 | AY301031 Zea mays |
| 23 | 16 | 88.9 | 365 | 8 | AY301037 Zea mays |
| 24 | 16 | 88.9 | 371 | 8 | AY301029 Zea mays |
| 25 | 16 | 88.9 | 371 | 8 | AY301032 Zea mays |
| 26 | 16 | 88.9 | 371 | 8 | AY301033 Zea mays |
| 27 | 16 | 88.9 | 372 | 8 | AY301035 Zea mays |
| 28 | 16 | 88.9 | 372 | 8 | AY301030 Zea mays |
| 29 | 16 | 88.9 | 384 | 8 | AY301034 Zea mays |
| 30 | 16 | 88.9 | 384 | 8 | AY301033 Zea mays |
| 31 | 16 | 88.9 | 559 | 8 | AY296473 Zea mays |
| 32 | 16 | 88.9 | 559 | 8 | AY296474 Zea mays |
| 33 | 16 | 88.9 | 559 | 8 | AY296476 Zea mays |
| 34 | 16 | 88.9 | 559 | 8 | AY296477 Zea mays |
| 35 | 16 | 88.9 | 559 | 8 | AY296481 Zea mays |
| 36 | 16 | 88.9 | 560 | 8 | AY296482 Zea mays |
| 37 | 16 | 88.9 | 561 | 8 | AY296458 Zea mays |
| 38 | 16 | 88.9 | 562 | 8 | AY296470 Zea mays |
| 39 | 16 | 88.9 | 562 | 8 | AY296475 Zea mays |
| 40 | 16 | 88.9 | 562 | 8 | AY296479 Zea mays |
| 41 | 16 | 88.9 | 562 | 8 | AY296480 Zea mays |
| 42 | 16 | 88.9 | 563 | 8 | AY296483 Zea mays |
| 43 | 16 | 88.9 | 563 | 8 | AY296456 Zea mays |
| 44 | 16 | 88.9 | 563 | 8 | AY296460 Zea mays |
| 45 | 16 | 88.9 | 563 | 8 | AY296462 Zea mays |
|    |    |      |     |   | AY296463 Zea mays |

## ALIGNMENTS

|                       |  |                                    |               |            |         |                 |
|-----------------------|--|------------------------------------|---------------|------------|---------|-----------------|
| RESULT 1              | AX342372   | Sequence 11 from Patent EP1167531. | 18 bp         | DNA        | linear  | PAT 12-JAN-2002 |
| LOCUS                 | AX342372   |                                    |               |            |         |                 |
| DEFINITION            | Sequence 11 from Patent EP1167531.   |                                    |               |            |         |                 |
| ACCESSION             | AX342372   |                                    |               |            |         |                 |
| VERSION               | AX342372.1   | GI:18151815                        |               |            |         |                 |
| KEYWORDS              |  |                                    |               |            |         |                 |
| SOURCE                |  |                                    |               |            |         |                 |
| ORGANISM              |  |                                    |               |            |         |                 |
| REFERENCE             |  |                                    |               |            |         |                 |
| AUTHORS               | Behr, C.F., Hironaka, C., Heck, G.R. and You, J.                                       |                                    |               |            |         |                 |
| TITLE                 | Corn transformant pv-zmgf32 (nke03) and compositions and methods for detection thereof |                                    |               |            |         |                 |
| JOURNAL               | Patent: EP 1167531-A 11 02-JAN-2002;   |                                    |               |            |         |                 |
| FEATURES              | Monsanto Technology LLC (US)   |                                    |               |            |         |                 |
| SOURCE                | Location/Qualifiers  |                                    |               |            |         |                 |
|                       | 1..18  |                                    |               |            |         |                 |
|                       | /organism="synthetic construct"  |                                    |               |            |         |                 |
|                       | /mol_type="unassigned DNA"   |                                    |               |            |         |                 |
|                       | /db_xref="taxon:32630"   |                                    |               |            |         |                 |
|                       | /note="Zea mays genomic DNA and vector DNA"  |                                    |               |            |         |                 |
| ORIGIN                |  |                                    |               |            |         |                 |
| Query Match           | 100.0%;  | Score 18;                          | DB 6;         | Length 18; |         |                 |
| Best Local Similarity | 100.0%;  | Pred. No. 5.8;                     |               |            |         |                 |
| Matches               | 18;  | Conservative 0;                    | Mismatches 0; | Indels 0;  | Gaps 0; |                 |
| Cy                    | 1 TGCCTGCTGCTGCTT 18   |                                    |               |            |         |                 |
| Db                    | 1 TGCCTGCTGCTGCTT 18   |                                    |               |            |         |                 |
| RESULT 2              | AX342369   | Sequence 8 from Patent EP1167531.  | 1183 bp       | DNA        | linear  | PAT 12-JAN-2002 |
| LOCUS                 | AX342369   |                                    |               |            |         |                 |
| DEFINITION            | Sequence 8 from Patent EP1167531.  |                                    |               |            |         |                 |
| ACCESSION             | AX342369   |                                    |               |            |         |                 |
| VERSION               | AX342369.1   | GI:18151812                        |               |            |         |                 |
| KEYWORDS              |  |                                    |               |            |         |                 |
| SOURCE                |  |                                    |               |            |         |                 |
| ORGANISM              |  |                                    |               |            |         |                 |
|                       | synthetic construct  |                                    |               |            |         |                 |
|                       | synthetic construct  |                                    |               |            |         |                 |

other sequences, artificial sequences.

1  
Behr, C.F., Hironaka, C., Heck, G.R. and You, J.  
Corn transformatant pv-zmg132 (nk603) and compositions and methods  
for detection thereof  
Patent: EP 1167531-A 8 02-JAN-2002;  
Monsanto Technology LLC (US)  
Location/Qualifiers

FEATURES  
source  
1. 1183  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="1-164 Agrobacterium tumefaciens nos 3' terminator  
165-381 construct vector DNA 382-686 Zea mays genomic DNA"  
genes, rps1 and rps4 687-1183 Zea mays genomic DNA"

ORIGIN  
Query Match 100.0%; Score 18; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
1 TGCTGTTCTGCTGACTTT 18  
|||||  
678 TGCTGTTCTGCTGACTTT 695

RESULT 3  
BC055684 2386 bp mRNA linear ROD 12-OCT-2004  
LOCUS  
Mus musculus zinc finger protein 445, mRNA (cDNA clone  
IMAGE:6410700), complete cds.  
BC055684  
BC055684.1 GI:33416952

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2386)

REFERENCE  
AUTHORS  
Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapelton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Kana, S.S., Loquellano, N.A., Peters, G.J.,  
Adams, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Sanchez, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2386)  
Director MGC Project.  
Direct Submission  
Submitted (01-AUG-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [gcgbs-remail.nih.gov](mailto:gcgbs-remail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-ehgc.stanford.edu>  
Contact: (Dickson, Mark) [md@pax1.stanford.edu](mailto:md@pax1.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAC Plate: 125 Row: 5 Column: 24  
This clone has the following problem: The cds is short compared to  
the longest cds in the locus.  
Location/Qualifiers

FEATURES  
source  
1. 2386  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6410700"  
/tissue\_type="Brain, enriched mouse brain 12.5dp"  
/clone\_lib="NIH BMAP\_F00"  
/lab\_host="DH10B"  
/note="Vector: pTX-ASC"  
1. 2386  
/gene="Zfp445"  
/note="synonym: ZNF168"  
/db\_xref="LOCUSID:235682"  
/db\_xref="WGI:2143340"  
319. 1482  
/gene="Zfp445"  
/codon\_start=1  
/product="Zfp445 protein"  
/protein\_id="AAH55684.1"  
/db\_xref="GI:33416953"  
/db\_xref="LOCUSID:235682"  
/db\_xref="WGI:2143340"  
/translation="MTLHNEKVKQKROEDFKQSYRQSGVISTVEKTEPCQNGKTF  
TOKSLIEHORITGEKPYQSCGCTFYRSYIIMRTQAIKKEHGLTFSQ  
GAVEPIPRGSHTEGSKYCGKAFHNSFLIHRVHTREKPYGRCCEKAFRFS  
NIVRQKQFLHRYKRYREKETSNSQXITLDKPPMOCGKFTFRRLDHDGI  
HSGERRKCNLCESKPDNRNRYLNHQRITTEDPQRDKDPVGHARSVDQRKSHVTL  
QSEGLHSDPRGSLCYCDVALNIOESGKURKCDNPSDESSKSIAPQVPTKQKCH  
KCSYCGTKFKXSHLISHKCHTKERPFCIVGKTFRWSNLTIRMKQHVNR"

ORIGIN  
Query Match 100.0%; Score 18; DB 10; Length 2386;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
1 TGCTGTTCTGCTGACTTT 18  
|||||  
1868 TGCTGTTCTGCTGACTTT 1885

RESULT 4  
BC034572 2699 bp mRNA linear ROD 12-OCT-2004  
LOCUS  
Mus musculus zinc finger protein 445, mRNA (cDNA clone  
IMAGE:4018895), partial cds.  
BC034572  
BC034572.1 GI:21961272

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2699)

REFERENCE  
AUTHORS  
Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,





## FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INLIL at: <http://image.lnl.gov>  
Series: IRAC Plate: 60 Row: 0 Column: 19.  
Location/Qualifiers

1. .2835

## gene

## CDS

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4953934"  
/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma, 5 month old virgin mouse."  
/clone\_1ib="NCI CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
1. .2835  
/gene="Zfp445"  
/note="synonym: ZNF168"  
/db\_xref="locusid:235682"  
/db\_xref="MGI:2143340"  
1. .1931  
/gene="Zfp445"  
/codon\_start=3  
/product="Zfp445 protein"  
/protein\_id="AA027167.1"  
/db\_xref="MGI:20071803"  
/db\_xref="locusid:235682"  
/db\_xref="MGI:2143340"  
/translation="RLQTSIGDSTIQMKEKAGDISORTGSESEVLNNDILKLVK  
CVSVSRKRLSPFKGYDRNPFKSSHHYNNKYGEGRGEGVYQNTGLKENGDRYG  
ETSRKSWHAPBHRQPSYSEGLFOCVGKAFKMSNRIRHKKIHGKPYOCSCLEKAFQ  
KAPORLSVRLHOKTHSKORSGSKYNALTCSDVSHLLTDRERKHLHNCCKGKAP  
SCGSVAIEHORIRHOKEPYKTRCRKTPFKMSNSRHKHLHKEVYQKROEPKPS  
YROSQVSTVEKTPPCNCGCTFOKSLIEHORIRHGERPYOCSCGERTYSSII  
IHMKRTIOHAIKIREHCLTFESQAVPIPIRGSHNTGSKNCKGAFNRSFLILH  
ERYVTEKPYOCSECEKAFKMSNLYRHORHFLKRYKESKETSNLQSKILIDK  
PFMOCEGKTFTRKSLIDHKGJHSGERFKCNLCESPDNRYLVNHORITTEPOWD  
WRDQDPVGIHARSVDORHNSLTLOSEYGLSDKGLSYCODVRLNIOELSGKLAKED  
NPDSKSLAFQVNPTRKKACHKSCSTCGKTFKSHLSHKKCHTTERPKCIVCGK  
TFRSSNLTRMKHVRN"

## ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 2835;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18

Db 2317 TGCTGTTCTGCTGACTTT 2334

RESULT 6  
AY341877 3788 bp mRNA linear ROD 19-AUG-2003  
LOCUS  
DEFINITION  
MUS musculus zinc finger protein 445 mRNA, complete cds.  
AY341877  
VERSION  
AY341877.1 GI:33638102  
KEYWORDS  
SOURCE  
MUS musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 3788).  
Zhou, G., Wang, J. and Zhang, Y.  
Cloning of mouse zinc finger protein 445  
Unpublished  
2 (bases 1 to 3788)  
Zhou, G., Wang, J. and Zhang, Y.  
Direct Submission  
Submitted (14-JUL-2003) Dept. of Genetics, Institute of Genetics,  
No 220, Handan Rd., Shanghai, Shanghai 200433, P.R.China  
Location/Qualifiers

## source

## CDS

1. .3788  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/Chromosome="9"  
66. .3026  
/codon\_start=1  
/product="zinc finger protein 445"  
/protein\_id="AA024161.1"  
/db\_xref="GI:33638103"  
/translation="MPPGRWHARSAQVREBOGCLRMVKEEDGYISQIARPQTLN  
RPOEIRFQLRQLRHESGPLETISRLQELCRMWRPDIKSKOMLELVLQFLS  
ILGELRTWYQLCPESGAELVALLLELDLDLDTLPCLQNDVDMITGSLAFQ  
AOWSPASHKNSALEDEHETSHGIGCDVLAEOGDSAVSPDYFOLEGEIEQEA  
LTPDYEVTFSEBOGCLNSAONLYRDLVINYGVAVSVSPRALISMLBARKP  
WGNICTVQLKRDADAAPREGKIQITPNKFIKQKSEIIEACVTSVSPKSVSEET  
GLKSEFKQSRIGTSCGDSIQMKEKAGDISORTGSESEVLNNDILKLVKCVS  
SRRLSPFKGYDRNPFKSSHHYNNKYGEGRGEGVYQNTGLKENGDRYG  
KSNHAPBHRQPSYSEGLFOCVGKAFKMSNRIRHKKIHGKPYOCSCLEKAFQ  
KAPORLSVRLHOKTHSKORSGSKYNALTCSDVSHLLTDRERKHLHNCCKGKAP  
VYIEHORIRHOKEPYKTRCRKTPFKMSNSRHKHLHKEVYQKROEPKPSYROS  
QVSTVEKTPPCNCGCTFOKSLIEHORIRHGERPYOCSCGERTYSSIIHMK  
RTQHALIKIREHCLTFESQAVPIPIRGSHNTGSKNCKGAFNRSFLILHKEVH  
TRKPYOCSECEKAFKMSNLYRHORHFLKRYKESKETSNLQSKILIDQEPWC  
OECGKTFTRKSLIDHKGJHSGERFKCNLCESPDNRYLVNHORITTEPOWD  
DFVGIHARSVDORHNSLTLOSEYGLSDKGLSYCODVRLNIOELSGKLAKED  
ESSKSLAFQVNPTRKKACHKSCSTCGKTFKSHLSHKKCHTTERPKCIVCGK  
TFRSSNLTRMKHVRN"

## ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 3788;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18

Db 3411 TGCTGTTCTGCTGACTTT 3428

RESULT 7  
AC125374 175440 bp DNA linear ROD 25-NOV-2003  
LOCUS  
DEFINITION  
MUS musculus BAC clone RP23-438P19 from chromosome 9, complete  
sequence.  
AC125374  
VERSION  
AC125374.4 GI:30089835  
KEYWORDS  
HTG.  
SOURCE  
MUS musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 175440).  
Dignan, G. and Haglund, K.  
The sequence of Mus musculus BAC clone RP23-438P19  
Unpublished (2001)  
2 (bases 1 to 175440)  
WILSON, R.  
TITLE  
Sequencing of Mus musculus  
JOURNAL  
Unpublished (2001)  
REFERENCE  
3 (bases 1 to 175440)  
McPherson, J.D. and Waterston, R.H.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (24-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 175440)  
McPherson, J.D. and Waterston, R.H.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (03-APR-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 175440)  
McPherson, J.D. and Waterston, R.H.  
TITLE  
Direct Submission

## JOURNAL

Submitted (24-APR-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA

## REFERENCE

6 (bases 1 to 175440)

## AUTHORS

Wilson, R.

## TITLE

Direct Submission

## JOURNAL

Submitted (25-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Apr 24, 2003 this sequence version replaced gi:29501931.  
Genome Center

## COMMENT

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: [submissions@genome.wustl.edu](mailto:submissions@genome.wustl.edu)

Center project name: M\_BA0438P19

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

SOURCE INFORMATION:  
The RPCI-23 BAC Library has been constructed by Kazutoyo Oseegawa  
and Minako Tateo in the laboratory of Pieter de Jong  
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or  
brain genomic DNA. The clone and detailed information can be  
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or  
Pieter de Jong and coworkers at <http://www.chori.org>

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

## FEATURES

## source

1..175440  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="9"  
/map="9"  
/clone="RP23-438P19"  
/clone\_lib="RPCI-23"  
repeat\_region  
/rpt\_family="Alu"  
637..781  
repeat\_region  
/rpt\_family="B2"  
871..1011  
repeat\_region  
/rpt\_family="B2"  
1421..1533  
repeat\_region  
/rpt\_family="B4"  
1813..1948  
repeat\_region  
/rpt\_family="B4"  
1949..2002  
repeat\_region  
/rpt\_family="B2"  
2123..2286  
repeat\_region  
/rpt\_family="Alu"  
2305..2481  
repeat\_region  
/rpt\_family="MER2\_type"  
3024..3138  
repeat\_region  
/rpt\_family="L1"  
3143..3318  
repeat\_region  
/rpt\_family="L1"

repeat\_region 3603..3736  
/rpt\_family="Alu"  
repeat\_region 4222..4324  
/rpt\_family="L1"  
repeat\_region 4343..4428  
/rpt\_family="B4"  
repeat\_region 4742..4914  
/rpt\_family="B2"  
repeat\_region 5197..5282  
/rpt\_family="MER2\_type"  
repeat\_region 5423..5493  
/rpt\_family="MER2\_type"  
repeat\_region 7280..7429  
/rpt\_family="Alu"  
repeat\_region 7815..7964  
/rpt\_family="Alu"  
repeat\_region 8235..8348  
/rpt\_family="MIR"  
repeat\_region 8485..8621  
/rpt\_family="Alu"  
repeat\_region 8706..8770  
/rpt\_family="ID"  
repeat\_region 9283..9524  
/rpt\_family="L1"  
repeat\_region 9575..9914  
/rpt\_family="MaLR"  
repeat\_region 9992..10122  
/rpt\_family="Alu"  
repeat\_region 10127..10223  
/rpt\_family="U6"  
repeat\_region 10453..10584  
/rpt\_family="Alu"  
repeat\_region 10602..10689  
/rpt\_family="Alu"  
repeat\_region 11339..11456  
/rpt\_family="Alu"  
repeat\_region 12314..12451  
/rpt\_family="Alu"  
repeat\_region 12620..12770  
/rpt\_family="Alu"  
repeat\_region 13406..13586  
/rpt\_family="B2"  
repeat\_region 13598..13647  
/rpt\_family="L2"  
repeat\_region 13697..13891  
/rpt\_family="ERVX"  
repeat\_region 14688..14992  
/rpt\_family="RMR12"  
repeat\_region 15621..16001  
/rpt\_family="MaLR"  
repeat\_region 16163..16348  
/rpt\_family="RMR12"  
repeat\_region 16944..17521  
/rpt\_family="ERV1"  
repeat\_region 18760..18903  
/rpt\_family="Alu"  
repeat\_region 19583..19723  
/rpt\_family="Alu"  
repeat\_region 19764..19870  
/rpt\_family="B4"  
repeat\_region 20295..20771  
/rpt\_family="RMR6B"  
repeat\_region 20804..20976  
/rpt\_family="ERVX"  
repeat\_region 20974..21254  
/rpt\_family="ERVX"  
repeat\_region 21255..21323  
/rpt\_family="ERVX"  
repeat\_region 21324..21840  
/rpt\_family="ERVX"  
repeat\_region 21803..21884  
/rpt\_family="ERV1"  
repeat\_region 21901..22284

```

repeat_region      /rpt_family="ERVK"
                    22726..22884
                    /rpt_family="B4"
                    22860..22923
                    /note="Sequence derived from one plasmid subclone."
unsure              22958..23159
                    /rpt_family="B2"
                    23398..24132
                    /rpt_family="Alu"
repeat_region      24532..24592
                    /rpt_family="B2"
repeat_region      24937..25150
                    /rpt_family="B4"
repeat_region      25176..25353
                    /rpt_family="B2"
repeat_region      25356..25438
                    /rpt_family="B4"
repeat_region      25914..26000
                    /rpt_family="S5"
repeat_region      28142..28283
                    /rpt_family="Alu"
repeat_region      28298..28352
                    /rpt_family="B2"

```

```

Query Match      100.0%; Score 18; DB 10; Length 175440;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCCTGTTCTGCTGACTTT 18
Db 105579 TGCCTGTTCTGCTGACTTT 105596

```

```

RESULT 8
AC094355 218326 bp DNA linear HTG 09-MAY-2003
LOCUS Rattus norvegicus clone CH230-3C12, *** SEQUENCING IN PROGRESS ***
DEFINITION 4 uncloned pieces.
ACCESSION AC094355 GI:30467252
VERSION AC094355.6
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

```

REFERENCE
AUTHORS 1 (bases 1 to 218326)
Muzny,D,Marle, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alibrooks,S, Amin,A, Anguiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Blawie,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Cesaar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Devila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Bacotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Gueraa,W,
Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
Hollins,B, Howells,S, Huylk,S, Hume,J, Idlebit,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpathy,S, Kelly,S, Khan,Z, King,L, Kovar,C,
Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorenzowa,L, Louised,H, Lozada,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindartne,M, Mahmood,M, Malloy,K, Mangum,A,

```

```

TITLE JOORNAL
REFERENCE JOORNAL
AUTHORS JOORNAL
TITLE JOORNAL
REFERENCE JOORNAL
AUTHORS JOORNAL
TITLE JOORNAL

```

## COMMENT

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced g1:24942540.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GALT

Center clone name: CH230-3C12

Summary Statistics

Assembly program: Atlas

Consensus quality: 199312 bases at least Q40

Consensus quality: 201330 bases at least Q30

Consensus quality: 202750 bases at least Q20

Estimated insert size: 207083; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 4 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 19397: contig of 19397 bp in length  
\* 19398 19497: gap of unknown length  
\* 19398 204378: contig of 184881 bp in length  
\* 204379 204478: gap of unknown length  
\* 204479 216833: contig of 12355 bp in length  
\* 216834 216933: gap of unknown length  
\* 216934 218326: contig of 1393 bp in length.

FEATURES  
source  
1. 218326 Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-3C12"  
204479..206130  
/note="wgs\_contig"

## ORIGIN

Query Match 100.0%; Score 18; DB 2; Length 218326;  
Best local Similarity 100.0%; Pred. No. 3.4;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 65609 TGCTGTTCTGCTGACTTT 65626

RESULT 9  
AC132091/c 140256 bp DNA linear ROD 30-JUL-2004  
LOCUS Mus musculus chromosome 18 clone RP24-46SH15, complete sequence.  
DEFINITION AC132091  
AC132091.3 GI:50839052  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS Wilson, R.K.  
TITLE The sequence of Mus musculus clone  
JOURNAL Unpublished  
REFERENCE  
AUTHORS McPherson, J.D. and Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (29-AUG-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA

REFERENCE  
AUTHORS Wilson, R.K.  
TITLE Direct Submission  
JOURNAL Submitted (06-SEP-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 140256)

REFERENCE  
AUTHORS Wilson, R.K.  
TITLE Direct Submission  
JOURNAL Submitted (30-JUL-2004) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On Jul 30, 2004 this sequence version replaced gi:34495013.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC

Web site: http://genome.wustl.edu  
Contact: submissions@wustl.edu  
----- Project Information -----  
Center project name: M\_BB0465H15  
-----

## FEATURES

source  
1. 140256  
/organism="Mus musculus"  
/mol\_type="genomic DNA"

/db\_xref="taxon:10090"  
/chromosome="18"  
/clone="RP24-46SH15"

## ORIGIN

Query Match 94.4%; Score 17; DB 10; Length 140256;  
Best local Similarity 100.0%; Pred. No. 14;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCTGTTCTGCTGACTTT 18  
|||||  
Db 69580 GCTGTTCTGCTGACTTT 69564

RESULT 10  
AC016256/c 164405 bp DNA linear PRI 25-JUL-2001  
LOCUS Homo sapiens 12 BAC RP11-98B6 (Roswell Park Cancer Institute Human  
BAC library) complete sequence.  
DEFINITION AC016256  
AC016256.19 GI:14670071  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 164405)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaralunga, H.C., Aze, J.R., Banks, T., Barbara, J.,  
Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,  
Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F.,  
Carter, M., Cavazos, S.R., Chacko, J., Chaver, D., Chen, G., Chen, R.,  
Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleveland, C.D.,  
Cox, C., Coyte, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,  
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D.,  
Flanagan, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J.,  
Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W.,  
Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K.,  
Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O.,  
Hodgson, A., Hogue, M., Hollway, C., Hollins, B., Homel, F.,  
Howard, S., Huber, J., Hulik, S., Hume, J., Ioshikhes, I., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C.,  
Lewie, L., Li, J., Li, Z., Lichteage, O., Lieu, C., Liu, J., Liu, W.,  
Loulsegad, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapa, F., Marondei, I., Martin, R.,  
Martindale, A., Martinez, E., Masey, E., Mashiney, E., McLeod, M.P.,  
Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G.,  
Miner, Z., Mitchell, T., Mohabhat, K., Montgomery, K.T., Morgan, M.,  
Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,  
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenkwu, S.,  
Ogun, M., Okwou, G., Oragunye, N., Oriole, R., Pace, A., Payton, B.,  
Peery, J., Perez, L., Peters, L., Pickens, R., Pitman, E., Pu, L.,  
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojuboakan, I., Rolfe, M.,  
Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shim, C.,  
Snooshari, N., Sison, I., Sodergren, E., Sonalke, T., Sparks, A.,  
Stanley, H., Stone, H., Sutton, A., Swatek, A., Tabor, P., Tamerisa, A.,  
Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telifrod, B.,  
Thomas, N., Thomas, S., Umanai, K., Vasquez, L., Vera, V., Villalon, D.,  
Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R.,  
Washington, C., Watlington, S., Williams, G., Williamson, A.,  
Wiczak, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J.,  
Zorrilla, S., Zuchelapatti, R. and Gibbs, R.

JOURNAL Submitted (24-NOV-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 164405)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 164405)

Worley,K.C.

REFERENCE Direct Submission

AUTHORS Submitted (25-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL On Jul 11, 2001 this sequence version replaced gi:12656764.

COMMENT INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES

source Location/Qualifiers

1..164405

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="12"

/clone="RP11-98B6"

/complement(1..371)

/rpt\_family="WLR1F1"

3..5166

/note="Overlaps bases 186760..191923 of clone AC087886"

/function="Overlaps with adjacent clone AC087886"

repeat\_region /rpt\_family="L1MA8"

2142..2281

/rpt\_family="MIR"

repeat\_region /complement(2827..2961)

/rpt\_family="MIR"

3943..4314

repeat\_region

/rpt\_family="L2"

repeat\_region 5502..5527

/rpt\_family="AT\_rich"

repeat\_region 6069..6615

/rpt\_family="L1ME3A"

repeat\_region 6830..6911

/rpt\_family="L2"

repeat\_region 7015..7068

/rpt\_family="AT\_rich"

repeat\_region 7069..7153

/rpt\_family="(TA)n"

repeat\_region 7163..7191

/rpt\_family="(TG)n"

repeat\_region complement(7192..7324)

/rpt\_family="L2"

repeat\_region complement(7387..7463)

/rpt\_family="L2"

repeat\_region 7825..8933

/rpt\_family="L1PA13"

repeat\_region complement(8937..9171)

/rpt\_family="L2"

repeat\_region 9494..9629

/rpt\_family="L1"

repeat\_region complement(9700..10075)

/rpt\_family="L1MA9"

repeat\_region 10071..10234

/rpt\_family="L1MA9"

repeat\_region 10294..10581

/rpt\_family="L1MA4"

repeat\_region 10582..11009

/rpt\_family="Tigger2a"

repeat\_region 11010..11303

/rpt\_family="L1M4"

repeat\_region 11500..11670

/rpt\_family="L1M3"

repeat\_region 11701..12100

/rpt\_family="L1M3"

repeat\_region 12185..13027

/rpt\_family="L1M3"

repeat\_region complement(13224..13505)

/rpt\_family="L1PA10"

repeat\_region 13829..13957

/rpt\_family="AT\_rich"

repeat\_region complement(13958..15549)

/rpt\_family="L1PA3"

14218

/note="A or G"

/function="Polymorphic site"

14264

/note="C or A"

/function="Polymorphic site"

14283

/note="C or A"

/function="Polymorphic site"

14288

/note="C or T"

/function="Polymorphic site"

14300..14431

/standard\_name="G19948"

/db\_xref="dbSTS:32826"

14356

/note="C or T"

/function="Polymorphic site"

14362

/note="C or G"

/function="Polymorphic site"

14397..14398

/note="CA or TG"

/function="Polymorphic site"

14426

misc\_feature

Query Match 94.4%; Score 17; DB 9; Length 164405;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 TCGTGTTCGTCGACTT 17  
 |||||  
 Db 51082 TCGTGTTCGTCGACTT 51066

RESULT 11  
 LMP696 172148 bp DNA linear INV 08-AUG-2001  
 LOCUS Leishmania major Friedlin chromosome 14 PAC P696, PREFINAL.  
 ACCESSION AL512293  
 VERSION AL512293.2 GI:15142995  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

1 (bases 1 to 172148)  
 Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and Smith,D.F.  
 A physical map of the Leishmania major Friedlin genome  
 Genome Res. 8 (2), 135-145 (1998)  
 9477341  
 2 (bases 1 to 172148)  
 Masny,D., Purnelle,B., Goffeau,A., Ivens,A.C., Quail,M., Rajadream,M.A. and Barrell,B.G.  
 Direct Submission  
 Submitted (07-AUG-2001) European Leishmania major Friedlin genome  
 sequencing project, Sanger Centre, The Wellcome Trust Genome  
 Camps, Hinxton, Cambridge CB10 1SA, (E-mail: barrell@sanger.ac.uk)  
 and Unite de Biochimie physiologique, Universite Catholique de  
 Louvain, Place Croix du Sud 2-20, B-1348 Louvain-la-Neuve, Belgium  
 On Aug 9, 2001 this sequence version replaced gi:11967804.  
 see http://www.ebi.ac.uk/parasites/leish.html

Notes:  
 Details of leishmania sequencing at the Sanger Centre are available  
 on the World Wide Web.  
 see http://www.sanger.ac.uk/Projects/L\_major/  
 CDS are numbered using the following system eg P696.01. P696

(cosmid name), .01 (first CDS)  
 To make the PAC library Leishmania major Friedlin DNA was partially  
 digested with NotI prior to cloning into the NotI site of the PAC  
 vector pCYPAC2N.  
 The more significant matches with motifs in the PROSITE database  
 are also included but some of these may be fortuitous. The length  
 in codons is given for each CDS.  
 Usually the highest scoring match found by fasta -o is given for  
 CDS which show significant similarity to other CDS in the database.  
 Gene prediction is done using:

(1)  
 the FramePlot program of Bibb et al.,  
 Gene 30:157-166(1984) as implemented  
 at http://www.nih.gov/jp/~jun/cgi-bin/frameplot.pl. (2)  
 codon preference based on the codon usage table for Leishmania at  
 http://www.kazusa.or.jp/codon/  
 (3)

the Hexamer program which was written by Richard Durbin as an  
 integral part of the AKEB-based analysis tools for the C.elegans  
 Genome Sequencing Project. The program calculates the  
 log-likelihood score for a given DNA segment based on the frequency  
 of 6-mers, normalised for the base-pair composition of the genome.  
 The program was trained using a fasta file of confirmed leishmania  
 major coding sequences (CDS), i.e. from ATG start codon to the stop  
 codon.

CAUTION: We may not have predicted the correct initiation codon.  
 where possible we choose an initiation codon (atg) which is  
 preceded by a stretch of pyrimidines or part of a Kozak sequence.  
 If this cannot be identified we choose the most upstream initiation  
 codon. Transmembrane domains were predicted as implemented at the  
 TMHMM server: http://www.cbs.dtu.dk/services/TMHMM-1.0/  
 IMPORTANT: This sequence MAY NOT be the entire insert of the

sequenced clone. It may be shorter because we only sequence  
 overlapping sections once, or longer, because we arrange for a  
 small overlap between neighbouring submissions.

#### FEATURES

source  
 1.172148  
 /organism="Leishmania major"  
 /mol\_type="genomic DNA"  
 /strain="Friedlin"  
 /db\_xref="taxon:5664"  
 /chromosome="14"  
 /clone="PAC P696"

#### ORIGIN

Query Match 94.4%; Score 17; DB 3; Length 172148;  
 Best Local Similarity 100.0%; Pred.No.13;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 TCGTGTTCGTCGACTT 17  
 |||||  
 Db 87315 TCGTGTTCGTCGACTT 87299

RESULT 12  
 AC139156 175062 bp DNA linear ROD 27-NOV-2003  
 LOCUS Mus musculus BAC clone RP24-115J20 from chromosome 18, complete  
 DEFINITION  
 ACCESSION AC139156  
 VERSION AC139156.3 GI:33342438  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

1 (bases 1 to 175062)  
 Shah,N.  
 The sequence of Mus musculus BAC clone RP24-115J20  
 Unpublished (2001)  
 2 (bases 1 to 175062)  
 Wilson,R.  
 Sequencing of Mus musculus  
 Unpublished (2001)  
 3 (bases 1 to 175062)  
 McPherson,J.D. and Waterston,R.H.  
 Direct Submission  
 Submitted (24-JAN-2003) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 4 (bases 1 to 175062)  
 Wilson,R.K.  
 Direct Submission  
 Submitted (26-JUN-2003) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 5 (bases 1 to 175062)  
 Wilson,R.K.  
 Direct Submission  
 Submitted (30-JUL-2003) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 6 (bases 1 to 175062)  
 Wilson,R.  
 Direct Submission  
 Submitted (27-NOV-2003) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Jul 30, 2003 this sequence version replaced gi:32261442.

COMMENT  
 JOURNAL

Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@wustl.edu  
 Summary Statistics  
 Center project name: M\_BB0115J20

NOTICE: This sequence may not represent the entire insert of this



clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

**SOURCE INFORMATION:**  
The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

**NEIGHBORING SEQUENCE INFORMATION:**  
This sequence is the entire insert of the clone. This clone is overlapped by AC131796 and AC117191.

#### FEATURES

##### source

1.175062  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="18"  
/map="18"  
/clone="RP24-115J20"  
/clone\_id="RPCI-24"  
2816.2832  
/rpt\_family="Alu"  
2833.3199  
/rpt\_family="ERVK"  
3200.3270  
/rpt\_family="Alu"  
3284.4049  
/rpt\_family="L1"  
4050.4441  
/rpt\_family="ERV1"  
4589.4801  
/rpt\_family="L1"  
4804.5251  
/rpt\_family="L1"  
5267.6350  
/rpt\_family="L1"  
6369.6443  
/rpt\_family="ERVK"  
6607.6651  
/rpt\_family="ERVK"  
7020.8857  
/rpt\_family="ERVK"  
9542.9763  
/rpt\_family="ERVK"  
9900.10153  
/rpt\_family="ERVK"  
10154.10574  
/rpt\_family="ERVK"  
10577.11619  
/rpt\_family="L1"  
13136.13296  
/rpt\_family="L1"  
13297.13771  
/rpt\_family="ERV1"  
13772.14140  
/rpt\_family="L1"  
15017.15329

/rpt\_family="L1"  
18263.18690  
/rpt\_family="ERVK"  
19452.19708  
/rpt\_family="ERVK"  
20128.20850  
/rpt\_family="ERVK"  
21320.21462  
/rpt\_family="ERV1"  
22527.22679  
/rpt\_family="ERV1"  
22867.23274  
/rpt\_family="ERVK"  
23506.23612  
/rpt\_family="ERVK"  
23814.24356  
/rpt\_family="L1"  
24451.25827  
/rpt\_family="L1"  
25898.26118  
/rpt\_family="MALR"  
28719.29029  
/rpt\_family="L1"  
29051.30039  
/rpt\_family="L1"  
30145.30382  
/rpt\_family="L1"  
31301.31495  
/rpt\_family="B4"  
31879.32154  
/rpt\_family="ERV1"  
32431.32622  
/rpt\_family="B2"  
33168.33241  
/rpt\_family="B2"  
33781.33952  
/rpt\_family="B2"  
34262.34352  
/rpt\_family="Alu"  
34418.34886  
/rpt\_family="L1"  
41760.42428  
/rpt\_family="L1"  
42404.42444  
/rpt\_family="L1"  
44749.44822  
/rpt\_family="MALR"  
45324.45314  
/rpt\_family="B4"  
45830.46019  
/rpt\_family="B2"  
46075.46440  
/rpt\_family="MALR"  
46551.46750  
/rpt\_family="B2"  
46787.46896  
/rpt\_family="B4"  
47633.47812  
/rpt\_family="B2"  
47972.48062  
/rpt\_family="L1"  
48820.49034  
/rpt\_family="B4"  
49400.49484  
/rpt\_family="L1"  
49700.49854  
/rpt\_family="Alu"  
49948.50004  
/rpt\_family="L1"  
50087.50185  
/rpt\_family="Alu"  
50130.50186  
/rpt\_family="B4"



```

repeat_region      50322..50451
                    /rpt_family="MairR"
repeat_region      51460..51517
                    /rpt_family="L1"
repeat_region      51610..51885

Query Match      94.4%; Score 17; DB 10; Length 175062;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GCGTCTCGCTGACTTT 18
Db      59764 GCGTCTCGCTGACTTT 59780

RESULT 13
AC105823/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-141P22, WORKING DRAFT SEQUENCE, 7
unordered pieces.
AC105823      229391 bp      DNA      linear      HTG 15-NOV-2002
AC105823
AC105823.4 GI:25007047
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229391)
Wozny,D,Marie, Metzker,M,Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandarenka,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blych,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacero,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falle,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,J., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Healand,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,D., Jollivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,Z., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenzowa,L., Louissege,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Mallory,K., Mangun,A.,
Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Mloasivjevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwokolemeah,O., Okwunonu,G., Olariunsaagun,A., Pal,S., Parks,K.,
Paternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Polidexter,A., Popovic,D., Primus,E., Pu,L., L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojars,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvatsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sned,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabors,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemari,K.,
Valdes,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNVH
Center clone name: CH230-141P22
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 209637 bases at least Q40
Consensus quality: 212405 bases at least Q30
Consensus quality: 213820 bases at least Q20
Estimated insert size: 21551; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 10067: contig of 10067 bp in length
* 10068 10167: gap of unknown length
* 10168 30107: contig of 19940 bp in length
* 30108 30207: gap of unknown length
* 30208 219981: contig of 189774 bp in length
* 219981 220081: gap of unknown length
* 220081 224729: contig of 4648 bp in length
* 224729 224829: gap of unknown length
* 224829 225855: contig of 1026 bp in length
* 225855 225955: gap of unknown length
* 225955 227516: contig of 1561 bp in length
* 227516 227517: gap of unknown length
* 227517 229391: contig of 1775 bp in length.
*
* Location/Qualifiers
* 1..229391

```

FEATURES  
Source

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-141P22"
misc_feature
    1..1175
    /note="wgs_contig"
misc_feature
    8319..11067
    /note="wgs_contig"
misc_feature
    10168..11538
    /note="wgs_contig"
misc_feature
    22082..221295
    /note="wgs_contig"

ORIGIN
Query Match      94.4% Score 17; DB 2; Length 229391;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGGTGTTCCTGACTT 17
Db      117028 TGGTGTTCCTGACTT 117012

RESULT 14
AC142509      231074 bp   DNA      linear   HTG 03-APR-2003
LOCUS        Rattus norvegicus clone CH230-unknown, WORKING DRAFT SEQUENCE, 38
DEFINITION   unordered pieces.
AC142509
VERSION      AC142509.1 GI:29501816
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM     Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 231074)
Muzny,D,Marie, Metzker,M,Lee, Abramzon,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Bunay,C, Burch,P, Butrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Dublin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Gaur,M,
Gheorgheghe,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevar,W,
Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
Hollins,B, Howells,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpethy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Koyar,Z,
Kowis,C, Kraft,C,L, Lebow,H, Levan,U, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorenshuwa,L, Louisedge,H, Lozada,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,
Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Mawhinney,S, McLeod,M, McNeill,T, Meenen,E, Milosavljevic,A,
Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,W, Morris,K,
Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D,
Newton,N, Nguyen,N, Norris,S, Nwaokemele,O, Okwona,G,
Olapunsaogun,A, Pal,S, Parks,K, Paaternak,S, Paul,H,
Perez,A, Perez,L, Pfankoch,C, Plopper,F, Poldexter,A,
Popovic,D, Primus,E, Pu,L-L, Puzo,M, Quiroz,J, Rachlin,E,
Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y,
Reuter,M, Richards,S, Riggs,F, River,C, Rodkey,T, Rojase,A,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
Scott,G., Shateman,S., Shen,H., Shetty,J., Shwartzbeyn,A.,
Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A.,
Swatek,A., Tabore,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Umani,K., Valae,R., Vera,V., Villaana,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center Project name: KZZY
Center clone name: CH230-unknown
-----
Summary Statistics
Sequencing vector: Plasmid,
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 216536 bases at least Q40
Consensus quality: 221871 bases at least Q30
Consensus quality: 225261 bases at least Q20
Estimated insert size: 214970, sum-of-ctnigs estimation
Quality coverage: 5x in Q20 bases; sum-of-ctnigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 38 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1537: contig of 1537 bp in length
1538 1637: gap of unknown length
1638 2663: contig of 1026 bp in length
2664 2763: gap of unknown length
2764 3850: contig of 1087 bp in length
3851 3950: gap of unknown length
3951 5198: contig of 1248 bp in length
5199 5298: gap of unknown length
5299 6650: contig of 1352 bp in length
6651 6750: gap of unknown length
6751 7852: contig of 1102 bp in length
7853 7952: gap of unknown length
7953 9555: contig of 1603 bp in length
9556 9656: gap of unknown length
9657 11097: contig of 1442 bp in length
11098 11197: gap of unknown length
11198 12750: contig of 1553 bp in length
12751 12850: gap of unknown length
12851 14233: contig of 1383 bp in length
14234 14333: gap of unknown length
14334 15578: contig of 1245 bp in length
15579 15679: gap of unknown length
15680 16807: contig of 1129 bp in length
16808 16907: gap of unknown length
16908 18834: contig of 1927 bp in length

```

[illegible]

VERSIONS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

HG097799.7 GI:30521368  
 HTGS\_PASSEL; HTGS\_DAAFT; HTGS\_FULLLTOP.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 231140)  
 Wuzny,D,Marle, Metzker,M,Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D, Anyalaebchi,V, Aoyagi,A, Ayodeji,M, Baca,E, Bader,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bismato,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedertich,D, Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Dublin,K, Duval,T, Evans,K, Egan,A, Becotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,J, Foster,M, Foster,P, Friszer,C,M, Gabisi,A, Gante,R, Garcia,A, Garner,T, Garza,W, Gegeorgis,E, Geer,K, Giller,R, Grady,M, Guerra,M, Guevara,W, Gunaratne,P, Haaland,W, Hamli,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Huily,S, Hume,J, Idlebird,J, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Kapathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorensuewa,L, Louisedge,H, Lozado,R,J, Lu,X, Ma,J, Maheswari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangun,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mahoney,S, Mcleod,M,P, McNeill,T,Z, Meenen,E, Milosevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Parks,K, Nwokilehem,O, Okunou,G, Olampungagun,A, Pal,S, Parks,K, Paternack,S, Paul,H, Perez,A, Perez,L, Pfannkoch,C, Plopper,F, Poidexter,A, Popovic,D, Primus,B, Pu,L, Pu,P, Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Reiger,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Slison,I, Sitter,C,D, Smaj,D, Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J, Steidle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umanet,K, Valas,R, Vera,V, Villasaana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wlecyk,R, Woden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,T, Zhou,X, Zhao,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstein,G, and Gibbs,R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 231140)  
 Worley,K.C.  
 Direct Submission  
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 231140)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On May 10, 2003 this sequence version replaced gi:24819446.  
 The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

#### ----- Genome Center

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

#### ----- Project Information

Center project name: GFGG  
Center clone name: CH230-46A18

#### ----- Summary Statistics

Assembly program: Atlas 3.0;  
Consensus quality: 218455 bases at least Q40  
Consensus quality: 220324 bases at least Q30  
Consensus quality: 222143 bases at least Q20  
Estimated insert size: 231276; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length.  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 228435: contig of 228436 bp in length  
\* 228437 228536: gap of unknown length  
\* 228537 231140: contig of 2604 bp in length.

#### FEATURES

##### source

1. 231140

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-46A18"

##### misc\_feature

1. 1042

/note="wgs\_end\_extension"

##### misc\_feature

clone\_end:T7"

5228. 8197

/note="wgs\_end\_extension"

##### misc\_feature

clone\_end:T7"

/note="clone\_boundary"

clone\_end:T7"

site:EcORI

end\_sequence:BH328184"

##### misc\_feature

complement(227335..228129)

/note="clone\_boundary"

clone\_end:Sp6

site:EcORI

end\_sequence:BH328185"

#### ORIGIN

Query Match 94.4%; Score 17; DB 2; Length 231140;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTT 17

DB 34562 TGCTGTTCTGCTGACTT 34546

Search completed: February 10, 2005, 13:13:54  
Job time : 114.027 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 3366.47 Seconds  
(without alignments)  
5630.828 Million cell updates/sec

Title: US-10-790-430-7

Perfect score: 498  
Sequence: 1 aatcgcattccaataatcgcgac.....aaaagtcgccaagttaa 498

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hlc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_g981.\*  
9: gb\_g982.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 67    | 13.5        | 670    | 6  | CB867772 HC01D06w  |
| 2          | 50    | 10.0        | 598    | 6  | CB862324 HH05GL2Y  |
| 3          | 34    | 6.8         | 793    | 9  | CL786064 OR_BBA010 |
| 4          | 29    | 5.8         | 539    | 8  | CC046788 3591.1.19 |
| 5          | 26    | 5.2         | 733    | 9  | CL858930 OR_CBA009 |
| 6          | 25    | 5.0         | 893    | 9  | CL795021 OR_CBA000 |
| 7          | 23    | 4.6         | 576    | 8  | AO689332 nbxb0079G |
| 8          | 23    | 4.6         | 617    | 8  | AO255572 nbxb0011D |
| 9          | 23    | 4.6         | 654    | 8  | CL615769 OR_BBA000 |
| 10         | 23    | 4.6         | 745    | 8  | AO326905 nbxb0039P |
| 11         | 23    | 4.6         | 778    | 8  | AO689711 nbxb0080M |
| 12         | 23    | 4.6         | 850    | 8  | AO573760 nbxb0084F |
| 13         | 23    | 4.6         | 854    | 8  | CM504264 OP_Ba001  |
| 14         | 22    | 4.4         | 801    | 9  | CM523527 OP_Ba004  |
| 15         | 22    | 4.4         | 702    | 9  | CL812779 OR_CBA002 |
| 16         | 22    | 4.4         | 713    | 9  | CG002331 ZUADU30TV |
| 17         | 22    | 4.4         | 733    | 9  | CL858930 OR_CBA009 |
| 18         | 22    | 4.4         | 784    | 9  | CL720122 OR_BBA004 |
| 19         | 22    | 4.4         | 891    | 9  | CL811054 OR_CBA002 |
| 20         | 22    | 4.4         | 893    | 9  | CL795021 OR_CBA000 |
| 21         | 21    | 4.2         | 654    | 9  | CL615769 OR_BBA000 |
| 22         | 21    | 4.2         | 680    | 9  | CL194957 104.419.1 |
| 23         | 21    | 4.2         | 692    | 8  | AZ600338 1M0418C05 |
| 24         | 21    | 4.2         | 702    | 9  | CL812779 OR_CBA002 |

|    |    |     |      |   |                    |
|----|----|-----|------|---|--------------------|
| 25 | 21 | 4.2 | 738  | 9 | CL832021 OR_CBA005 |
| 26 | 21 | 4.2 | 746  | 8 | AO870229 nbe0036N  |
| 27 | 21 | 4.2 | 786  | 9 | AG475192 Mus muscu |
| 28 | 21 | 4.2 | 801  | 9 | CL811054 OR_CBA002 |
| 29 | 21 | 4.2 | 811  | 8 | CC018239           |
| 30 | 21 | 4.2 | 854  | 8 | CC004665           |
| 31 | 21 | 4.2 | 1348 | 2 | BE967273 601661142 |
| 32 | 20 | 4.0 | 276  | 8 | BZ245511 CH230-314 |
| 33 | 20 | 4.0 | 352  | 8 | CC383351 PUBIO15TD |
| 34 | 20 | 4.0 | 386  | 9 | CC702094 OCUF058TH |
| 35 | 20 | 4.0 | 452  | 9 | AG206444 Oryza sat |
| 36 | 20 | 4.0 | 587  | 8 | AO921459 RPI1-23-2 |
| 37 | 20 | 4.0 | 620  | 8 | BH11311 RPI1-24-3  |
| 38 | 20 | 4.0 | 645  | 8 | CG065648 PUJBR28TB |
| 39 | 20 | 4.0 | 657  | 8 | BZ126584 CH230-490 |
| 40 | 20 | 4.0 | 657  | 9 | CG913725 ZMMBBD037 |
| 41 | 20 | 4.0 | 661  | 9 | CG065651 PUJBR28TD |
| 42 | 20 | 4.0 | 686  | 2 | BF274572 GA_E002   |
| 43 | 20 | 4.0 | 699  | 9 | CR155347 Reverse 8 |
| 44 | 20 | 4.0 | 720  | 9 | AG575889 Mus muscu |
| 45 | 20 | 4.0 | 730  | 5 | BM932647 UI-M-CG0P |

#### ALIGNMENTS

RESULT 1  
CB867772  
LOCUS  
DEFINITION  
CB867772 CH Hordeum vulgare cDNA clone HC01D06 5-PRIME, mRNA  
sequence.

ACCESSION  
CB867772  
VERSION  
CB867772.1 GI:30069748

KEYWORDS  
SOURCE  
ORGANISM

EST.  
Hordeum vulgare  
Hordeum vulgare  
Bukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
1 (bases 1 to 670)

REFERENCE  
AUTHORS  
Varshney, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and  
Graner, A.

TITLE  
JOURNAL  
COMMENT  
Barley ESTs from coleoptile tissue  
Unpublished (2003)  
Contact: Stein Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de

Insert Length: 670 Std Error: 0.00  
Plate: 1 row: D column: 6  
Seq primer: T7.  
Location/Qualifiers

FEATURES  
source

1..670  
/organism="Hordeum vulgare"  
/mol\_type="mRNA"  
/cuiivar="Sloop"  
/db\_xref="GABI:511307"  
/db\_xref="taxon:4513"  
/clone="HC01D06"  
/tissue\_type="coleoptile"  
/dev\_stage="coleoptile, 3 days old"  
/lab\_host="DH10B"  
/clone\_lib="CH"  
/note="Vector: pSPOT; Site 1: SalI (5-end of cDNA);  
Site 2: NotI (3-end of cDNA). Due to the cloning system  
used blue/white selection for recombinants is not 100%  
reliable. Average insert size is 1.3 kb."

ORIGIN  
Query Match 13.5%; Score 67; DB 6; Length 670;

Best Local Similarity 100.0%; Pred. No. 4.3e-26;  
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAAACAAGCTAA 418  
|||||  
DB 58 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAAACAAGCTAA 117  
|||||

QY 419 GATTACC 425  
|||||  
DB 118 GATTACC 124  
|||||

RESULT 2  
LOCUS CB862324 598 bp mRNA linear EST 22-APR-2003  
DEFINITION HH05G12Y HH Hordeum vulgare cDNA clone HH05G12 3-PRIME, mRNA  
sequence.

ACCESSION CB862324  
VERSION CB862324 GI:30056883  
KEYWORDS EST.  
SOURCE Hordeum vulgare  
ORGANISM Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
1 (bases 1 to 598)  
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and  
Graner,A.

REFERENCE  
AUTHORS Graner,A.  
TITLE Barley ESTs from coleoptile tissue  
JOURNAL Unpublished (2003)  
COMMENT Contact: Stein Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de  
Insert Length: 598 Std Error: 0.00  
Plate: 5 row: G column: 12  
Seq primer: SP6.  
Location/Qualifiers  
1..598  
/organism="Hordeum vulgare"  
/mol\_type="mRNA"  
/culivar="Sloop"  
/db\_xref="GABI:533740"  
/db\_xref="taxon:4513"  
/clone="HH05G12"  
/issue\_type="coleoptile"  
/dev\_stage="coleoptile, 1 day old"  
/lab\_host="HDH10B"  
/clone\_lib="HH"  
/note="Vector: pSPORT; Site\_1: SalI (5-end of cDNA);  
Site\_2: NciI (3-end of cDNA); Due to the cloning system  
used blue/white selection for recombinants is not 100 %  
reliable. Average insert size is 1.3 kb."

FEATURES  
source  
Location/Qualifiers  
1..598  
/organism="Hordeum vulgare"  
/mol\_type="mRNA"  
/culivar="Sloop"  
/db\_xref="GABI:533740"  
/db\_xref="taxon:4513"  
/clone="HH05G12"  
/issue\_type="coleoptile"  
/dev\_stage="coleoptile, 1 day old"  
/lab\_host="HDH10B"  
/clone\_lib="HH"  
/note="Vector: pSPORT; Site\_1: SalI (5-end of cDNA);  
Site\_2: NciI (3-end of cDNA); Due to the cloning system  
used blue/white selection for recombinants is not 100 %  
reliable. Average insert size is 1.3 kb."

ORIGIN  
Query Match 10.0%; Score 50; DB 6; Length 598;  
Best Local Similarity 100.0%; Pred. No. 2e-16;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAA 408  
|||||  
DB 21 TCGAGGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATAA 70  
|||||

RESULT 3  
LOCUS CL786064 793 bp DNA linear GSS 30-JUL-2004  
DEFINITION OR\_BB80103K12.f OR\_BBa Oryza rufipogon genomic clone OR\_BB80103K12  
5', genomic survey sequence.  
ACCESSION CL786064

VERSION CL786064.1 GI:50865715  
KEYWORDS GSS.  
SOURCE Oryza rufipogon  
ORGANISM Oryza rufipogon  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 793)  
Klim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,  
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.

REFERENCE  
AUTHORS Klim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,  
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.  
TITLE OMAP Project  
JOURNAL Unpublished (2004)  
COMMENT Contact: Rod A. Wing  
Arizona Genomics Institute  
University of Arizona  
Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA GG  
Insert Length: 161 Std Error: 0.00  
Plate: 0103 row: K column: 12  
Seq primer: TAA TAC GAC TCA CTA TAG GG  
Class: BAC ends.  
Location/Qualifiers  
1..793  
/organism="Oryza rufipogon"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4529"  
/clone="OR\_BB80103K12"  
/issue\_type="young leaves"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OR\_BBa"  
/note="Vector: pGIRAC1; Site\_1: HindIII; Site\_2: HindIII"

ORIGIN  
Query Match 6.8%; Score 34; DB 9; Length 793;  
Best Local Similarity 100.0%; Pred. No. 2.3e-07;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGT 389  
|||||  
DB 720 TACTCGAGTCATTCATATGCTTGAGAGAGAGT 753  
|||||

RESULT 4  
LOCUS CC046788/c 539 bp DNA linear GSS 01-APR-2003  
DEFINITION 3591.1 190.1 F04.Y 1 3591 - RescueMu Grid P Zea mays genomic,  
genomic survey sequence.  
ACCESSION CC046788  
VERSION CC046788.1 GI:29461679  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
Clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 539)  
Walbot,V.  
Maize genomic sequences found using engineered RescueMu transposon  
unpublished (2001)  
Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 723 8221  
Email: walbot@stanford.edu  
Plate: 3591\_1 190\_1 row: 33  
Class: transposon-tagged.

FEATURES  
source  
Location/Qualifiers  
1. .539  
/organism="zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/873/K55"  
/db\_xref="taxon:4577"  
/issue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_id="3591 - Rescuedu Grid P"  
/note="Organ: Leaf; Vector: Rescuedu (engineered from pluscript backbone); Site 1: BamHI; Site 2: BglII; Rescuedu is a 4.9 kb, modified maize Ma transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuedu, go to the web site 'www.zmud.iasrate.edu' and follow the links for 'Rescuedu', Grid P was grown at Molokai in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN  
Query Match 5.8%; Score 29; DB 8; Length 539;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 CTGAGACGTGCGTCCCTGCTGGAGCTGCT 281  
|||||  
443 CTGAGACGTGCGTCCCTGCTGGAGCTGCT 415

RESULT 5  
LOCUS CL858930 733 bp DNA linear GSS 16-AUG-2004  
DEFINITION OR\_CBA0090A04.r OR\_CBA Oryza rufipogon genomic clone OR\_CBA0090A04  
ACCESSION CL858930  
VERSION CL858930.1 GI:51270762  
KEYWORDS GSS.  
SOURCE Oryza rufipogon  
ORGANISM Oryza rufipogon  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Bharatoidae; Oryzaceae; Oryza.  
1 (bases 1 to 733)  
Kim,H., Yu,Y., Wisotsaki,M., Yost,D., Stum,D., Rao,K., Luo,M.,  
Jelty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and  
Wing,R.  
TITLE OMAP project  
JOURNAL Unpublished (2004)  
COMMENT Contact: Rod A. Wing  
Arizona Genomics Institute  
University of Arizona  
Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Plate: 0090 row: A column: 04  
Seq primer: CAC TCA TTA GGC ACC CCA  
Class: BAC ends.  
Location/Qualifiers  
1. .733  
/organism="Oryza rufipogon"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4529"  
/clone="OR\_CBA0090A04"  
/issue\_type="young leaves"  
/dev\_stage="2 week old seedlings"  
/dev\_stage="2 week old seedlings"

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

ORIGIN  
Query Match 5.2%; Score 26; DB 9; Length 893;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 ATCGTAATAAAGTGCCCAAGT 495  
|||||  
365 ATCGTAATAAAGTGCCCAAGT 340

RESULT 6  
LOCUS CL795021 893 bp DNA linear GSS 06-AUG-2004  
DEFINITION OR\_CBA004L03.r OR\_CBA Oryza rufipogon genomic clone OR\_CBA004L03  
ACCESSION CL795021  
VERSION CL795021.1 GI:51017037  
KEYWORDS GSS.  
SOURCE Oryza rufipogon  
ORGANISM Oryza rufipogon  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Bharatoidae; Oryzaceae; Oryza.  
1 (bases 1 to 893)  
Kim,H., Yu,Y., Wisotsaki,M., Yost,D., Stum,D., Rao,K., Luo,M.,  
Jelty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and  
Wing,R.  
TITLE OMAP project  
JOURNAL Unpublished (2004)  
COMMENT Contact: Rod A. Wing  
Arizona Genomics Institute  
University of Arizona  
Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Plate: 0004 row: L column: 03  
Seq primer: CAC TCA TTA GGC ACC CCA  
Class: BAC ends.  
Location/Qualifiers  
1. .893  
/organism="Oryza rufipogon"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4529"  
/clone="OR\_CBA004L03"  
/issue\_type="young leaves"  
/dev\_stage="2 week old seedlings"  
/lab\_host="DH10B T1 phage resistant"  
/clone\_id="OR\_CBA"  
/note="Vector: PACIBAC1; Site 1: HindIII; Site 2: HindIII;  
dkf treated 36 hrs before harvest"

ORIGIN  
Query Match 5.2%; Score 26; DB 9; Length 893;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 ATCGTAATAAAGTGCCCAAGT 495  
|||||  
365 ATCGTAATAAAGTGCCCAAGT 340

RESULT 7  
LOCUS AO689332 576 bp DNA linear GSS 01-JUL-1999

DEFINITION nbxb0079612f CUGI Rice BAC library Oryza sativa (japonica cultivar-group) genomic clone nbxb0079612f, genomic survey

ACCESSION AO689332

VERSION AO689332.1 GI:5330500

KEYWORDS GSS.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1 (bases 1 to 576)  
Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998)

AUTHORS Wing, R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu

Seq primer: TAATACGACTCACAATAGCG  
Class: BAC ends  
High quality sequence stop: 231.  
Location/Qualifiers

FEATURES

source

1. 576  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="nbxb0079612f"  
/issue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice BAC library"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 5.0%; Score 25; DB 8; Length 576;  
Best local Similarity 100.0%; Pred. No. 0.028;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 474 GTATATAAGGTGGCCCAAGTGA 498  
|||||  
Db 93 GTATATAAGGTGGCCCAAGTGA 69

RESULT 8  
AO255572 617 bp DNA linear GSS 23-OCT-1998  
LOCUS nbxb0011D02r CUGI Rice BAC library Oryza sativa (japonica  
DEFINITION cultivar-group) genomic clone nbxb0011D02r, genomic survey  
sequence.  
ACCESSION AO255572  
VERSION AO255572.1 GI:3780054

KEYWORDS GSS.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1 (bases 1 to 617)  
Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998)

AUTHORS Wing, R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu

Seq primer: GGAACACGCTATACCATG  
Class: BAC ends  
High quality sequence start: 11  
High quality sequence stop: 296.  
Location/Qualifiers

FEATURES

source

1. 617  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="nbxb0011D02r"  
/issue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice BAC library"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 617;  
Best local Similarity 100.0%; Pred. No. 0.38;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 371 ATATGCTTGAGAGAGAGTGGG 393  
|||||  
Db 309 ATATGCTTGAGAGAGAGTGGG 311

RESULT 9  
CL615769 654 bp DNA linear GSS 01-JUL-2004  
LOCUS OR\_BBA0008D16.r OR\_BBA Oryza rufipogon genomic clone OR\_BBA0008D16  
DEFINITION 3', genomic survey sequence.  
ACCESSION CL615769  
VERSION CL615769.1 GI:49104999  
KEYWORDS GSS.  
SOURCE Oryza rufipogon  
ORGANISM Oryza rufipogon  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Ehrhartoideae; Oryzae; Oryza.  
1 (bases 1 to 654)  
Kim, H., Yu, Y., Stum, D., Yost, D., Rao, K., Luo, M., Jettly, R.,  
Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.  
OMAP Project  
Unpublished (2004)  
Contact: Rod A. Wing  
Arizona Genomics Institute  
University of Arizona  
Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: htcp://genome.arizona.edu  
PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Insert length: 161 Std Error: 0.00  
Plate: 0008 row: D column: 16  
Seq primer: CAC TCA TTA GGC ACC CCA  
Class: BAC ends.  
location/Qualifiers  
1..654  
/organism="Oryza rufipogon"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4529"  
/clone="OR\_BBa0008D16"  
/issue\_type="young leaves"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OR\_BBa"  
/note="Vector: pNG1BAC1, Site\_1: HindIII; Site\_2: HindIII"

ORIGIN

Query Match 4.6%; Score 23; DB 9; Length 654;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATATGCTTGAGAGAGAGTCGGG 393  
|||||  
Db 311 ATATGCTTGAGAGAGAGTCGGG 333

RESULT 10  
AQ326905 745 bp DNA linear GSS 08-JAN-1999  
LOCUS nbxb0039P05r CUGI Rice BAC Library Oryza sativa (japonica  
DEFINITION cultivar-group) genomic clone nbxb0039P05r, genomic survey  
sequence.  
ACCESSION AQ326905.1 GI:4118755  
VERSION AQ326905  
KEYWORDS GSS.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
TITLE Ehrhartoideae; Oryzae; Oryza.  
JOURNAL 1 (bases 1 to 745)  
COMMENT 1 (bases 1 to 745)  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: twing@clemson.edu  
Seq primer: GGAAACAGCTATGACCAGT  
Class: BAC ends  
High quality sequence start: 4  
High quality sequence stop: 316.  
location/Qualifiers  
1..745  
/organism="Oryza sativa (japonica cultivar-group)"

FEATURES  
source

/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="nbxb0039P05r"  
/issue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice BAC Library"  
/note="Vector: pBeloBAC11, Site\_1: HindIII; Site\_2:  
HindIII; Rice is one of two most popular grains in the  
world. Half of the world population especially those  
inhabiting highly populated areas of the humid tropics  
and subtropics, rely on rice as their primary source of  
carbohydrate. Monocotyledonous rice is a diploid plant  
(2n=24) with a haploid genome equivalent of 431 Mbp  
(Arumuganathan and Earle, 1991). The relatively small  
genome of rice, three times larger than that of  
Arabidopsis, makes it suitable for genomic studies. In  
order to facilitate positional cloning, physical mapping  
and genome sequencing of rice, we have constructed a BAC  
library from Oryza sativa, Nipponbare variety. The  
library contains 36,864 clones with an average insert size  
of 128.5 Kb providing 10.9 haploid genome equivalents. The  
deep coverage allows the isolation a particular sequence  
with a probability of 99.9%. Two high density filters,  
each containing 18,432 clones (doubly spotted), represent  
the whole library for colony screening."

ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 745;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATATGCTTGAGAGAGTCGGG 393  
|||||  
Db 309 ATATGCTTGAGAGAGTCGGG 331

RESULT 11  
AQ689711 778 bp DNA linear GSS 01-JUL-1999  
LOCUS nbxb0080M19f CUGI Rice BAC Library Oryza sativa (japonica  
DEFINITION cultivar-group) genomic clone nbxb0080M19f, genomic survey  
sequence.  
ACCESSION AQ689711.1 GI:5330879  
VERSION AQ689711  
KEYWORDS GSS.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
TITLE Ehrhartoideae; Oryzae; Oryza.  
JOURNAL 1 (bases 1 to 778)  
COMMENT 1 (bases 1 to 778)  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: twing@clemson.edu  
Seq primer: TAAATGAGCTCATCTATGAGG  
Class: BAC ends  
High quality sequence stop: 279.  
location/Qualifiers  
1..778  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="nbxb0080M19f"  
/issue\_type="Leaf"

FEATURES  
source

/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice BAC Library"  
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:  
HindIII; Rice is one of two most popular grains in the  
world. Half of the world population especially those  
inhabiting highly populated areas of the humid tropics  
and subtropics, rely on rice as their primary source of  
carbohydrate. Monocotyledonous rice is a diploid plant  
(2n=24) with a haploid genome equivalent of 431 Mbp  
(Arumuganathan and Earle, 1991). The relatively small  
genome of rice, three times larger than that of  
Arabidopsis, makes it suitable for genomic studies. In  
order to facilitate positional cloning, physical mapping  
and genome sequencing of rice, we have constructed a BAC  
library from *Oryza sativa*, Nipponbare variety. The  
library contains 36,864 clones with an average insert size  
of 128.5 kb providing 10.9 haploid genome equivalents. The  
deep coverage allows the isolation a particular sequence  
with a probability of 99.9%. Two high density filters,  
each containing 18,432 clones (doubly spotted), represent  
the whole library for colony screening."

## ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 778;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 371 ATATGCTTGAGAGAGAGTCGGG 393

Db 309 ATATGCTTGAGAGAGAGTCGGG 331

## RESULT 12

AO573760 850 bp DNA linear GSS 02-JUN-1999  
LOCUS AO573760  
DEFINITION nbxb0084f24r CUGI Rice BAC Library *Oryza sativa* [japonica  
cultivar-group] genomic clone nbxb0084f24r, genomic survey  
sequence.

ACCESSION AO573760  
VERSION AO573760  
KEYWORDS GSS.  
SOURCE AO573760.1 GI:4974545

## ORGANISM

*Oryza sativa* (japonica cultivar-group)  
*Oryza sativa* (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; *Oryza*.

## REFERENCE

1 (bases 1 to 850)  
Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: GGAAGACGATATGACCATG  
Class: BAC ends  
High quality sequence stop: 310.  
Location/Qualifiers

## JOURNAL

COMMENT

## FEATURES

source

1. .850  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="nbxb0084f24r"  
/issue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice BAC Library"  
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:  
HindIII; Rice is one of two most popular grains in the  
world. Half of the world population especially those

inhabiting highly populated areas of the humid tropics  
and subtropics, rely on rice as their primary source of  
carbohydrate. Monocotyledonous rice is a diploid plant  
(2n=24) with a haploid genome equivalent of 431 Mbp  
(Arumuganathan and Earle, 1991). The relatively small  
genome of rice, three times larger than that of  
Arabidopsis, makes it suitable for genomic studies. In  
order to facilitate positional cloning, physical mapping  
and genome sequencing of rice, we have constructed a BAC  
library from *Oryza sativa*, Nipponbare variety. The  
library contains 36,864 clones with an average insert size  
of 128.5 kb providing 10.9 haploid genome equivalents. The  
deep coverage allows the isolation a particular sequence  
with a probability of 99.9%. Two high density filters,  
each containing 18,432 clones (doubly spotted), represent  
the whole library for colony screening."

## ORIGIN

Query Match 4.6%; Score 23; DB 8; Length 850;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 371 ATATGCTTGAGAGAGAGTCGGG 393

Db 309 ATATGCTTGAGAGAGAGTCGGG 331

## RESULT 13

CM504264/c 854 bp DNA linear GSS 06-OCT-2004  
LOCUS CM504264/c  
DEFINITION OP\_Ba0015P11.f OP\_Ba *Oryza punctata* genomic clone OP\_Ba0015P11  
5', genomic survey sequence.

ACCESSION CM504264  
VERSION CM504264  
KEYWORDS GSS.  
SOURCE CM504264.1 GI:53833769

## ORGANISM

*Oryza punctata*  
*Oryza punctata*  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; *Oryza*.

## REFERENCE

1 (bases 1 to 854)  
SanMiguel, P., Westerman, R., Kim, H., Yu, Y., Wisotski, M., Yost, D.,  
Stum, D., Rao, K., Luo, M., Jettly, R., Kudrna, D., Miller, C.,  
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S.A.  
OMAP Project - Purdue University  
Unpublished (2004)  
Contact: Scott A. Jackson  
Jackson Laboratory  
Purdue University  
915 W. State St., West Lafayette, IN 47907, USA  
Tel: 7654963621  
Fax: 7654967255  
Email: sjackson@purdue.edu

## JOURNAL

COMMENT

## FEATURES

source

1. .854  
/organism="Oryza punctata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4537"  
/clone="OP\_Ba0015P11"  
/issue\_type="young leaves"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OP\_Ba"

ORIGIN /note="Vector: PAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

Query Match 4.6%; Score 23; DB 9; Length 854;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 460 AAAGTAAATATCGTAAATAAA 482  
DB 350 AAAGTAAATATCGTAAATAAA 328

RESULT 14  
CM523527 601 bp DNA linear GSS 08-OCT-2004  
LOCUS OP\_Ba0049120.r OP\_Ba Oryza punctata genomic clone OP\_Ba0049120  
DEFINITION 3', genomic survey sequence.

ACCESSION CM523527 GI:53997749  
VERSION CM523527.1  
KEYWORDS GSS.  
SOURCE Oryza punctata  
ORGANISM Oryza punctata

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriarthroideae; Oryzaceae; Oryza.  
1 (bases 1 to 601)  
Samtuel, P., Westerman, R., Kim, H., Yu, Y., Misotski, M., Yost, D.,  
Stum, D., Rao, K., Luo, M., Jettly, R., Kudrna, D., Muller, C.,  
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S. A.  
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S. A.

TITLE OMAP Project - Purdue University  
JOURNAL Unpublished (2004)  
COMMENT Contact: Scott A. Jackson  
Jackson Laboratory  
Purdue University  
915 W. State St., West Lafayette, IN 47907, USA  
Tel: 765/4963621  
Fax: 765/4967255  
Email: sjackson@purdue.edu

Basecalling by phred version 0.020425.c. This sequence was derived  
from the raw sequence read by clipping with licy version 1.198.  
Bases 55-655 of the raw sequence (length 1219) were retained after  
clipping.

PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Insert length: 161000 Std Error: 0.00  
Plate: 0049 row: 1 column: 20  
Seq primer: CAC TCA TTA GGC ACC CCA  
Class: BAC ends.

FEATURES  
source location/Qualifiers

1..601  
/organism="Oryza punctata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4537"  
/clone="OP\_Ba0049120"  
/issue\_type="young leaves"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OP\_Ba"  
/note="Vector: PAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

ORIGIN

Query Match 4.4%; Score 22; DB 9; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 365 TCATTCAATGCTTGAGAAGAG 386  
DB 261 TCATTCAATGCTTGAGAAGAG 240

RESULT 15  
CL812779 702 bp DNA linear GSS 09-AUG-2004  
LOCUS

DEFINITION OR\_CBa0028M15.r OR\_CBa Oryza rufipogon genomic clone OR\_CBa0028M15  
3', genomic survey sequence.

ACCESSION CL812779 GI:51054117  
VERSION CL812779.1  
KEYWORDS GSS.  
SOURCE Oryza rufipogon  
ORGANISM Oryza rufipogon

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriarthroideae; Oryzaceae; Oryza.  
1 (bases 1 to 702)  
Kim, H., Yu, Y., Misotski, M., Yost, D., Stum, D., Rao, K., Luo, M.,  
Jettly, R., Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and  
Wing, R.

TITLE OMAP Project  
JOURNAL Unpublished (2004)  
COMMENT Contact: Rod A. Wing  
Arizona Genomics Institute  
University of Arizona  
Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Plate: 0028 row: M column: 15  
Seq primer: CAC TCA TTA GGC ACC CCA  
Class: BAC ends.

FEATURES  
source location/Qualifiers

1..702  
/organism="Oryza rufipogon"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4529"  
/clone="OR\_CBa0028M15"  
/issue\_type="young leaves"  
/dev\_stage="2 week old seedlings"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OR\_CBa"  
/note="Vector: PAGIBAC1; Site\_1: HindIII; Site\_2: HindIII;  
drk treated 36 hrs before harvest"

ORIGIN

Query Match 4.4%; Score 22; DB 9; Length 702;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 371 ATATGCTTGAGAAGAGTCGG 392  
DB 310 ATATGCTTGAGAAGAGTCGG 331

Search completed: February 10, 2005, 17:01:45  
Job time : 3370.47 secs

**This Page Blank (uspto)**





REFERENCE  
AUTHORS  
TITLE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 Ruf, M. and Koesel, H.  
Structure and expression of the gene coding for the alpha-subunit of DNA-dependent RNA polymerase from the chloroplast genome of Zea mays  
Nucleic Acids Res. 16 (13), 5741-5754 (1988)

FEATURES  
source

## CDS

/product="rps11 protein (1 is 2nd base in codon)"  
/protein\_id="CAA30659.1"  
/db\_xref="GI:829328"  
/db\_xref="GOA:P09561"  
/db\_xref="UniProt/Swiss-Prot:P09561"  
/translation="PPQYCGFKSKRASPAGQRTAVDAIRTVGLQRAEYVWKGAGSG RDAALRAIKSGVRLSCIRVTPMHPNGCPKRRRL"  
289..293  
/note="pot. RNA binding site"  
311..1336  
/codon\_start=1  
/transl\_table=1  
/product="RNA polymerase alpha-subunit (AA 1-341)"  
/protein\_id="CAA30670.1"  
/db\_xref="GI:12467"  
/db\_xref="GOA:P09562"  
/db\_xref="UniProt/Swiss-Prot:P09562"  
/translation="WVRELTGSTQLEMKVESRVSCKLYGRFLLSPKQADT VGLAARALLGEIGTCITRAKPGNVHEVSTTIGESIOELINKEIYLRNLG VRDASTVKGRRITADIIIPSVEDITDTPRIANRREPVDPIEHLQIRDAVHTE LRKNSQGSYPIDAVPMVNVNVSIFS CGNKGHEILFLIETNGKLPKGGIYS ASRNLIDLPFLHTBEGTSFEENKRLTPPLLTFRFTNLKNNKGLPUNCIPTD QLEPSRTYCNLKRAINPSALDLVLVTEBDLILINSGMEDGLIWDLEKHLPIDL PKNKFSL"

## ORIGIN

Query Match 20.5%; Score 242; DB 8; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 5.1e-129;  
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

445 TATATTGATCATGTAATCGTTATTTCTCTGGAAGCGGTTCAATTTTTTTAGCA 504  
446 TATATTGATCATGTAATCGTTATTTCTCTGGAAGCGGTTCAATTTTTTTAGCA 504  
301 TATATTGATCATGTAATCGTTATTTCTCTGGAAGCGGTTCAATTTTTTTAGCA 242  
505 CGCTTTTATAGAGGTGCAATCATGATGCGGATAGGTGTATACATCGGTATACAA 564  
506 CGCTTTTATAGAGGTGCAATCATGATGCGGATAGGTGTATACATCGGTATACAA 564  
241 CGCTTTTATAGAGGTGCAATCATGATGCGGATAGGTGTATACATCGGTATACAA 182  
565 CTAAACGTACCACTTTTGAAGATGCGTGTAAATGGGCACTCTCTCCGTATACACA 624  
566 CTAAACGTACCACTTTTGAAGATGCGTGTAAATGGGCACTCTCTCCGTATACACA 624  
181 CTAAACGTACCACTTTTGAAGATGCGTGTAAATGGGCACTCTCTCCGTATACACA 122  
625 CCTTTACATTAATCTGCTCGTGTGCAAAACCACTGTAGCAATGACATCTA 684  
626 CCTTTACATTAATCTGCTCGTGTGCAAAACCACTGTAGCAATGACATCTA 684  
121 CCTTTACATTAATCTGCTCGTGTGCAAAACCACTGTAGCAATGACATCTA 62

QY 685 CT 686  
DB 61 CT 60

RESULT 4  
ZMA86563  
LOCUS 140384 bp DNA circular PLN 09-SEP-2004  
DEFINITION Zea mays complete chloroplast genome.  
X86563.2  
ACCESSION GI:11990232  
VERSION 16S rRNA: 23S rRNA: 4.5S rRNA: 5S rRNA: ATPase; atp6 gene; atp1 gene; cema gene; clp gene; cytochrome b/f complex; cytochrome b6; cytochrome f; infA gene; initiation factor 1; junction LSC-IR; matK gene; NDH dehydrogenase; ndhB gene; ndhC gene; ndhD gene; ndhF gene; ndhJ gene; ndhK gene; petA gene; petB gene; petD gene; petE gene; psaA gene; psbA gene; psbB gene; psbC gene; psbD gene; psbJ gene; psbK gene; psbN gene; PSI P700 apoprotein A1; PSI P700 apoprotein A2; PSII component; PSII cytochrome b559; PSII D2 protein; PSII L protein; PSII K protein; PSII L protein; PSII protein; rbcL gene; ribosomal protein L14; ribosomal protein L16; ribosomal protein L23; ribosomal protein L20; ribosomal protein L22; ribosomal protein L23; ribosomal protein L32; ribosomal protein L33; ribosomal protein L36; ribosomal protein L31; ribosomal protein S12; ribosomal protein S14; ribosomal protein S15; ribosomal protein S16; ribosomal protein S18; ribosomal protein S19; ribosomal protein S2; ribosomal protein S3; ribosomal protein S4; ribosomal protein S7; ribosomal protein S8; ribosomal RNA; RNA polymerase; RNA-Asparagine; rpl14 gene; rpl16 gene; rpl2 gene; rpl20 gene; rpl22 gene; rpl23 gene; rpl32 gene; rpl33 gene; rpl36 gene; rpoA gene; rpoB gene; rpoC1 gene; rpoC2 gene; rps11 gene; rps12 gene; rps14 gene; rps15 gene; rps16; rps18 gene; rps19 gene; rps2 gene; rps3 gene; Rps4 gene; rps7 gene; rps8 gene; Rubisco; transfer RNA-Ala; transfer RNA-Arg; transfer RNA-Asp; transfer RNA-Cys; transfer RNA-Glu; transfer RNA-Gly; transfer RNA-His; transfer RNA-Ile; transfer RNA-Leu; transfer RNA-Lys; transfer RNA-Met; transfer RNA-Phe; transfer RNA-Pro; transfer RNA-Ser; transfer RNA-Thr; transfer RNA-Tyr; transfer RNA-Ty; transfer RNA-Val.

SOURCE  
ORGANISM

chloroplast Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
AUTHORS  
TITLE

1 Scrittmacher, G. and Koesel, H.  
Characterization of the mRNA transcripts of the maize, ribulose-1, 5-bisphosphate carboxylase, large subunit gene  
Plant Mol. Biol. 4, 169-179 (1985)

REFERENCE  
AUTHORS  
TITLE

2 Ertion, J.L.  
Characterization of the mRNA transcripts of the maize, ribulose-1, 5-bisphosphate carboxylase, large subunit gene  
Plant Mol. Biol. 4, 169-179 (1985)

REFERENCE  
AUTHORS  
TITLE

3 Fish, L.E., Kuck, U. and Bogorad, L.  
Two partially homologous adjacent light-inducible maize chloroplast genes encoding polypeptides of the P700 chlorophyll a-protein complex of photosystem I  
J. Biol. Chem. 260 (3), 1413-1421 (1985)

REFERENCE  
AUTHORS  
TITLE

4 Rock, C.D., Barkan, A. and Taylor, W.C.  
The maize plastid psbB-psbF-psbH-psbI gene cluster: spliced and unspliced psbB and psbI RNAs encode alternative products  
Curr. Genet. 12 (1), 69-77 (1987)

|           |   |    |
|-----------|---|----|
| REFERENCE | McLaughlin, W.B. and Larrima, I.M.                                      | 5  |
| AUTHORS   | The sequence of the maize plastid encoded rpl 22 locus                  |    |
| TITLE     | Nucleic Acids Res. 15 (10), 4356 (1987)                                 |    |
| JOURNAL   | 87231016  |    |
| PUBMED    | 3295778   |    |
| REFERENCE | Rodermel, S., Orlin, P. and Bogorad, L.                                 | 6  |
| AUTHORS   | The transcription termination region between two                        |    |
| TITLE     | convergently-transcribed photoregulated operons in the maize            |    |
| JOURNAL   | plastid chromosome contains rpl14, trnK (UCU) and a putative trnfm      |    |
| PUBMED    | pseudogene  |    |
| REFERENCE | Nucleic Acids Res. 15 (13), 5493 (1987)                                 |    |
| AUTHORS   | 87260027  |    |
| TITLE     | 3601681   |    |
| JOURNAL   | Giese, K., Subramanian, A.R., Larrima, I.M. and Bogorad, L.             | 7  |
| PUBMED    | Nucleotide sequence, promoter analysis, and linkage mapping of the      |    |
| REFERENCE | unusually organized operon encoding ribosomal proteins S7 and S12       |    |
| AUTHORS   | in maize chloroplast  |    |
| TITLE     | J. Biol. Chem. 262 (31), 15251-15255 (1987)                             |    |
| JOURNAL   | 88033114  |    |
| PUBMED    | 2822717   |    |
| REFERENCE | Schantz, R. and Bogorad, L.   | 8  |
| AUTHORS   | Maize chloroplast genes ndhB, ndhE, and psac. Sequences,                |    |
| TITLE     | transcripts and transcript pools  |    |
| JOURNAL   | Plant Mol. Biol. 11, 239-247 (1988)                                     | 9  |
| PUBMED    | Luken, J.H. and Bogorad, L.   |    |
| REFERENCE | Nucleotide sequence containing the maize chloroplast proline (UGG)      |    |
| AUTHORS   | and tryptophan (CCA) tRNA genes   |    |
| TITLE     | Nucleic Acids Res. 16 (11), 5192 (1988)                                 |    |
| JOURNAL   | 88262508  |    |
| PUBMED    | 3387224   |    |
| REFERENCE | Ruf, M. and Kossel, H.  | 10 |
| AUTHORS   | Structure and expression of the gene coding for the alpha-subunit       |    |
| TITLE     | of DNA-dependent RNA polymerase from the chloroplast genome of Zea      |    |
| JOURNAL   | mays  |    |
| PUBMED    | Nucleic Acids Res. 16 (13), 5741-5754 (1988)                            |    |
| REFERENCE | 88289331  |    |
| AUTHORS   | 3399379   |    |
| TITLE     | 11  |    |
| JOURNAL   | Bowman, C.M., Barker, R.F. and Dyer, T.A.                               |    |
| PUBMED    | In wheat ctDNA, segments of ribosomal protein genes are dispersed       |    |
| REFERENCE | repeats, probably conserved by nonreciprocal recombination              |    |
| AUTHORS   | Curr. Genet. 14 (2), 127-136 (1988)                                     |    |
| TITLE     | 89028843  |    |
| JOURNAL   | 3180271   |    |
| PUBMED    | 12  |    |
| REFERENCE | Steinmuller, K., Ley, A.C., Steimetz, A.A., Sayre, R.T. and Bogorad, L. |    |
| AUTHORS   | Characterization of the ndhC-pdbG-ORF157/159 operon of maize            |    |
| TITLE     | plastid DNA and of the cyanobacterium Synechocystis sp. PCC6803         |    |
| JOURNAL   | Mol. Gen. Genet. 216 (1), 60-69 (1989)                                  |    |
| PUBMED    | 89281491  |    |
| REFERENCE | 2499764   |    |
| AUTHORS   | 13  |    |
| TITLE     | Igloi, G.L., Meinke, A., Dory, I. and Kossel, H.                        |    |
| JOURNAL   | Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon:        |    |
| PUBMED    | comparison between the derived protein primary structures from          |    |
| REFERENCE | various organisms with respect to functional domains                    |    |
| AUTHORS   | Mol. Gen. Genet. 221 (3), 379-394 (1990)                                |    |
| TITLE     | 90340289  |    |
| JOURNAL   | 2381419   |    |
| PUBMED    | 14  |    |
| REFERENCE | Fitzky, B. and Subramanian, A.R.  |    |
| AUTHORS   | Nucleotide sequence and map positions of the duplicated gene for        |    |
| TITLE     | chloroplast ribosomal protein S15 in Zea mays (maize)                   |    |
| JOURNAL   | Nucleic Acids Res. 18 (11), 3407 (1990)                                 |    |
| PUBMED    | 90287730  |    |
| REFERENCE | 2129550   |    |
| AUTHORS   | 15  |    |

| Query Match           | 20.5%  | Score 242        | DB 8     | Length 140384 |
|-----------------------|--|------------------|----------|---------------|
| Best Local Similarity | 100.0%   | Pred. No. 9e-129 |          |               |
| Matches 242           | Conservative 0   | Mismatches 0     | Indels 0 | Gaps 0        |
| 445                   | TATATTGATCATTCATCGTTATTCCTTGAAAGCGGTTCAATTTTTTTACAG        | 504              |          |               |
| 77199                 | TATATTGATCATTCATCGTTATTCCTTGAAAGCGGTTCAATTTTTTTACAG        | 77255            |          |               |
| 505                   | CGTCTTTTTTGAAGAGTCGACATCCATTATGCGGCAATAGGTGTACATCGGTATACAA | 564              |          |               |



Db 77259 CGCTTTTATAGAGTGCACATCCATTATGGCGCATAGTGTATCATGGCGTATACAA 77318  
QY 565 CTTAACCGTACACCACTTTTATAGCAATGCTGTATATCGGCATCTCTCCGCTACACGCA 624  
Db 77319 CTTAACCGTACACCACTTTTATAGCAATGCTGTATATCGGCATCTCTCCGCTACACGCA 77378  
QY 625 CCTTTATACCACTTCTGTCTGTGCAAAACCACTGTAGATATAGCATCTACGCGCT 684  
Db 77379 CCTTTATACCACTTCTGTCTGTGCAAAACCACTGTAGATATAGCATCTACGCGCT 77438  
QY 685 CT 686  
Db 77439 CT 77440

RESULT 5  
AE009947 141182 bp DNA circular PLN 09-JUN-2004  
LOCUS Saccharum hybrid cultivar SP-80-3280 chloroplast, complete genome.  
AE009947  
VERSION AE009947.2 GI:50080438  
KEYWORDS chloroplast Saccharum hybrid cultivar SP-80-3280  
SOURCE Saccharum hybrid cultivar SP-80-3280  
ORGANISM Saccharum hybrid cultivar SP-80-3280  
REFERENCE 1 (bases 1 to 141182)  
AUTHORS Calsa, T. Jr., Carraro, D.M., Benatti, M.R., Barbosa, A.C.,  
Kitaljima, J.P., and Carter, H.  
TITLE Structural features and transcript editing analysis of sugarcane  
JOURNAL (Saccharum officinarum L.) chloroplast genome  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 141182)  
Calsa, T. Jr.  
TITLE Direct Submission  
JOURNAL Submitted (21-JUN-2002) Ciencias Biologicas, Escola Superior de  
Agricoltura Luiz de Queiroz - Universidade de Sao Paulo, Av. Padua  
Dias, 11 - Agronomia, Piracicaba, SP 13418-900, Brazil  
COMMENT On Jul 9, 2004 this sequence version replaced gi:48478610.  
FEATURES  
source location/Qualifiers  
1..141182  
/organism="Saccharum hybrid cultivar SP-80-3280"  
/organism="Saccharum hybrid cultivar SP-80-3280"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:193079"  
complement(1139..510)  
/gene="rpl14"  
/locus\_tag="PS001"  
complement(139..510)  
/gene="rpl14"  
/locus\_tag="PS001"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein L14"  
/protein\_id="AA14631.1"  
/db\_xref="GI:48478611"  
/translation="MIQPTLLVADNSGARKLNCIRVIGAAAGNORVARIQDVIIVAVI  
KDAVQWPLERSEVIRAVIVTRKEFKDDDIIRYDNNAAVVIDQKNPGRVFGA  
VAEIRLEINLYKIVSLAPEVL"  
complement(620..2109)  
/gene="rpl16"  
/locus\_tag="PS002"  
complement(join(620..1021,2101..2109))  
/gene="rpl16"  
/locus\_tag="PS002"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein L16"  
/protein\_id="AA14632.1"  
/db\_xref="GI:48478612"  
/translation="WLSPKRTRFRKHGRNMGKSGCRNMHICFGKYLQVLEPAMITA

gene  
complement(2254..2928)  
/gene="rps3"  
/locus\_tag="PS003"  
complement(2254..2928)  
/gene="rps3"  
/locus\_tag="PS003"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein S3"  
/protein\_id="AA14633.1"  
/db\_xref="GI:48478613"  
/translation="MGQKINPLGFRLLGTQNNHSPWPAQPKYNSGLQEDKKIRNCIK  
NYIOKRRKGSNPKMSDSSSEVITHTIEIOKEIDTIVIHIGFPLNLKKKGAIEELE  
KDIQKSVSNQRLINAIKEKVEPRQPNILAEVIAFOKKNRVSFRKAKKAIELTK  
ADIKGIQIOIAGLAKETARACIKKGLPLQITIRAKIDYCCYPIRTIYGLGVKIV  
IFVEER"  
complement(2988..3434)  
/gene="rpl22"  
/locus\_tag="PS004"  
complement(2988..3434)  
/gene="rpl22"  
/locus\_tag="PS004"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein L22"  
/protein\_id="AA14634.1"  
/db\_xref="GI:48478614"  
/translation="MTSFKLVKTPRIKKKKGLRKLRKVPTRDLLEFEEVFAQKRI  
HMSVFAQRLVDEIRMYEETMIINLMYRASYPILKLYSANAATYRQDKTN  
LFTLKAESVSTMTKKPRPARGRSISIKTWCNITIVINIVKSK"  
complement(3527..3808)  
/gene="rps19"  
/locus\_tag="PS005"  
complement(3527..3808)  
/gene="rps19"  
/locus\_tag="PS005"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein S19"  
/protein\_id="AA14635.1"  
/db\_xref="GI:48478615"  
/translation="MTRKTNPFVARRLAKIEKVNKKEKEIIVTSRASSILPAMV  
GHTIATNGKEHTPIYTTPMVGKGEFPTNHFISYESTROTYSRR"  
complement(3828..4052)  
/locus\_tag="PS007"  
complement(3828..4052)  
/locus\_tag="PS007"  
/note="orf75"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein 75"  
/protein\_id="AA14636.1"  
/db\_xref="GI:48478616"  
/translation="MENCIRGNCNNMAGSERELNPRMVDSGTALIHATSAPYPAKG  
ESLSIHYISITSDIHTSIRIVDIGCHSLK"  
3939..4013  
/gene="trnH"  
/locus\_tag="PS006"  
3939..4013  
/gene="trnH"  
/locus\_tag="PS006"  
/product="trnA-His"  
4175..4588  
/locus\_tag="PS009"  
4175..4588  
/locus\_tag="PS009"  
/note="orf137"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein 137"  
/protein\_id="AA14637.1"



gene /cranal\_table=11  
/product="ribosomal protein S12"  
/protein\_id="BAD27316.1"  
/db\_xref="GI:49659535"  
/translation="MPTVKQLINRNAROPINNAKSAALKGCPORGTCARVYTINPK  
PNSLRKRVARVLTSGEITAYIPGIGHNIQHSVVLVIRGGRVVDLPGVAYRIIRGTL  
DAVAVNRQGRSKYGAKKPKK"  
complement (join(1389, .1423,3902, .3938))  
/gene="trnK"  
/complement (join(1389, .1423,3902, .3938))  
/gene="trnK"  
/product="tRNA-Lys"  
/note="codon recognized: AAA"  
complement (1673, .3274)  
/gene="matK"  
complement (1673, .3274)  
/gene="matK"  
/codon\_start=1  
/cranal\_table=11  
/product="matKase"  
/protein\_id="BAD27271.1"  
/db\_xref="GI:49659490"  
/translation="MYHLLINREMLFSLIOVEIOMKFESEYSEKOKSRQHYFVYPLLF  
OEYIVAFHNYNGSEPVPIPGCNKKPSSILVGRLLIRMYOONFLINSVYNODR  
LPHRNYFYSEFYSQIISSEFGIVVEIPLSLGQSCPEKEIPKQWLQSIHISIPFL  
EDKFLHLYSHIEIPYIHLEILVQLERYIQDVPSELHLIRFFLHYISNNGLITSM  
KSIPLKKNQQLFRFLNYSVSEFFFLFLRKQSSCLRLTSSGTLEHIIISGKME  
HFGVMYGFPRKTIWFPMDDLHYVRVQGAIIASKQTLILKKKMSYLVNFSQYFY  
FMTPORIRLNOLNISCDFLYLSVPIITLIVRNQMLNSFLIDTRMKKFTYVA  
TPIVGSLSKAOCFSGHPISKIPVMTDLSMDILDRFGRIICNLFIHSGSSKKQTIY  
RLKTIYRLSCARTLARKKSTVRFMRQLSVLEEFTEBEQVFSLMFTKTIHFSFH  
GSOSERIWYLDIIRINDLVNPLTN"  
complement (join(4503, .4720,5555, .5594))  
/gene="trpL6"  
complement (join(4503, .4720,5555, .5594))  
/gene="trpL6"  
/codon\_start=1  
/cranal\_table=11  
/product="ribosomal protein S16"  
/protein\_id="BAD27272.1"  
/db\_xref="GI:49659491"  
/translation="MVKLRKRCGRKQOATYRIVAIDVRSRREGDRLKVGFDPIKN  
QTCIANVPAIIYFELEKGAQPRTYVDIIRKAEFFDKERTIS"  
complement (7124, .7195)  
/gene="trnQ"  
complement (7124, .7195)  
/gene="trnQ"  
/product="tRNA-Gln"  
/note="codon recognized: CAA"  
7550, .7735  
/gene="psbK"  
7550, .7735  
/gene="psbK"  
/codon\_start=1  
/cranal\_table=11  
/product="PSII K-protein"  
/protein\_id="BAD27273.1"  
/db\_xref="GI:49659492"  
/translation="MPNLTSLTICPNSVLCPTSPFAKLEAAYAIFNPVIDVNPVLP  
VLPFLAIFWQAANSFR"  
8124, .8234  
/gene="psbI"  
8124, .8234  
/gene="psbI"  
/codon\_start=1  
/cranal\_table=11  
/product="PSII I-protein"  
/protein\_id="BAD27274.1"  
/db\_xref="GI:49659493"  
/translation="MLTLKLFYTVVIFVSLFIIFGLSNDPGRNPGREE"  
complement (8361, .8448)  
/gene="trnS"  
complement (8361, .8448)  
CDS

gene /gene="trnS"  
/product="tRNA-Ser"  
/note="codon recognized: AGC"  
9014, .9145  
/gene="orf43"  
9014, .9145  
/gene="orf43"  
/codon\_start=1  
/cranal\_table=11  
/product="hypothetical protein"  
/protein\_id="BAD27275.1"  
/db\_xref="GI:49659494"  
/translation="MNKGFOKSIPLFNIFLFFPRILLFPHPCNRRMRKXGFLKK"  
9425, .10486  
/gene="psbD"  
9425, .10486  
/gene="psbD"  
/codon\_start=1  
/cranal\_table=11  
/product="PSII D2 protein"  
/protein\_id="BAD27276.1"  
/db\_xref="GI:49659495"  
/translation="MTIAGRVYKEENDLFDIMDWLRDRFVFGWGCILLFPICAY  
ALGQFETGTFVTSWYTHGLASSYLEGCNFLTAAVSTPANSLSHSLLLMGPAQGF  
TRWCGGLMTFVALGAPALIGFMLOFELASVOLRPYNALISFGPIVFSVFLI  
YVIGSGWFPAPSGVAATPRFLLPFGGFNMTLNPFHMMGVAGVGAALLCAIHGAT  
VENTLPEDDGDGNTPRAPNPQAEFTYSMTYANRFSQIFGVAFSNRMLHFMFLFVP  
VTLNMSAIGVGLANLRYDVSOEIRAAEDBPETFTTKIILNEGIRAMMAAD  
QPHENLIIPREVLPRGNAL"  
10434, .11855  
/gene="psbC"  
10434, .11855  
/gene="psbC"  
/codon\_start=1  
/cranal\_table=11  
/product="PSII 44kD protein"  
/protein\_id="BAD27277.1"  
/db\_xref="GI:49659496"  
/translation="MKTIYSLRRFYVEVTLFNGTFVLACGRDOETTFPAMNAGNRLIN  
LSKLLGAHVAAHGLIVFMGAPNLPEVAHFVBEKPYEBOGLIILPHIALVLCGVGCG  
GEVLDTFPFYPSGVHLIISAVVIGFGLIYHALGPETLESFPFGVWVQDRNMTTI  
LGJHLILGLGAPLLVLCALYFGVGDVMAFGGDVARKIINLTLSPGVIFGYLLKSP  
GGGMIYVDDLDIIGCHVAGSI CVLGGIWHILTFPAMARBAFWSGSAVLSYSL  
AASVSGFICACFVFNMTAYPSEFGTGPBASQAQFFFLVRDOLGANVGAQCP  
TGIGKTIYMSPTGEVIFGSETMRFPDLRAPMLEIRGPNGLDLSRLKCDIQWQERS  
AEYMTAPLSGLSVGVAITEINAVNTVSPRSWLATSHFVLGFFPFVGHIMHGRABA  
AAAGFEKSIDRDLEPVLXMTPLN"  
complement (12010, .12097)  
/gene="trnS"  
complement (12010, .12097)  
/gene="trnS"  
/product="tRNA-Ser"  
/note="codon recognized: UCA"  
12441, .12629  
/gene="psbZ"  
12441, .12629  
/gene="psbZ"  
/codon\_start=1  
/cranal\_table=11  
/product="PSII Z-protein"  
/protein\_id="BAD27278.1"  
/db\_xref="GI:49659497"  
/translation="MTIAFQAVFALITSSVLVISVPLVPASPDGSSNNKNVFSGT  
SLWIGLVFVAIINSLIS"  
12915, .12985  
/gene="trnG"  
12915, .12985  
/gene="trnG"  
/product="tRNA-Gly"  
/note="codon recognized: GGC"  
complement (13477, .13550)  
/gene="trnfm"  
complement (13477, .13550)  
CDS

gene  
CDS  
gene  
CDS  
gene  
CDS



QY 163 TTTTATATATGTGAAAAGAGTAATTTCACTTGGGCCACCTTTTATTACCGATATTTT 222  
| | | | |  
Db 920 TTTTATATATGTGAAAAGAGTAATTTCACTTGGGCCACCTTTTATTACCGATATTTT 861  
| | | | |  
QY 223 ACTTTATACCACTTTTAACTGATGTTTCACTTTTGACCAAGGTAATCTTACCTTTGTTT 282  
| | | | |  
Db 860 ACTTTATACCACTTTTAACTGATGTTTCACTTTTGACCAAGGTAATCTTACCTTTGTTT 801  
| | | | |  
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 333  
| | | | |  
Db 800 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 750  
| | | | |  
RESULT 10  
AY452735/c 4895 bp DNA circular SYN 19-DEC-2003  
LOCUS Reporter vector pActXN, complete sequence.  
DEFINITION AY452735  
ACCESSION AY452735 GI:39636979  
VERSION  
KEYWORDS Reporter vector pActXN  
SOURCE Reporter vector pActXN  
ORGANISM other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 4895)  
AUTHORS Vickers,C.E., Xue,G.P. and Greshoff,P.M.  
TITLE A synthetic xylanase as a novel reporter in plants  
JOURNAL Plant Cell Rep. 22 (2), 135-140 (2003)  
MEDLINE 22867549  
PUBMED 12845475  
REFERENCE 2 (bases 1 to 4895)  
AUTHORS Vickers,C.E.  
TITLE Direct Submission  
JOURNAL Submitted (29-OCT-2003) ARC Centre for Integrative Legume Research,  
The University of Queensland, Room 213, John Hines Building (69),  
St. Lucia, QLD 4072, Australia  
FEATURES  
source  
1. .4895  
/organism="Reporter vector pActXN"  
/mol\_type="other DNA"  
/db\_xref="taxon:255329"  
267..2274  
misc\_feature  
/note="sequence from cloning vector pBR322"  
468..1328  
CDS  
/note="ampicillin resistance protein; bla"  
/codon\_start=1  
/product="beta-lactamase"  
/protein\_id="AAR29085.1"  
/db\_xref="GI:39636980"  
/translacion="MSIOHFRVALIPFPAAFCLEVFAPHPETLVKVKDAEDQLGARVGY  
IETDLSGKILIESFRPERFPMWSTFVLLCGAVLSRIDAQEQQLGRRIHYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAIIAMTSNDTANLLITIGDKELTAEFLHNMGDVTRL  
DRPEPELNEAIPIRDERPTTPVMATTLRKILTGSELITLSRQQLIDMEADKVAGPL  
LRSLAPGMFLPADSGAGERSRGIITAAIGDPKRSRIIVITTTGSAQTMDERRRQIA  
ETGASLIKHM"  
2345..2367  
misc\_feature  
/note="7 RNA polymerase promoter"  
2426..3338  
promoter  
/note="from rice actin gene; Act1"  
3339..3792  
intron  
/note="intron 1 from rice actin gene; Act1"  
3800..4513  
gene  
/gene="exYNA"  
3800..4513  
/gene="exYNA"  
/note="synthetic xylanase"  
/codon\_start=1  
/product="xylanase"  
/protein\_id="AAR29086.1"  
/db\_xref="GI:39636981"  
/translacion="WASNGKFTVGNQONHKNVNDGSEYETLNDTNGSGSMITLGS  
ATPKAEMAAVAVNRGFLARGLDPSOKKADYDYGIVATYATYKOTASAGNSLGS  
YGFQNRGLNGVPLVEYYITIEDMDVMDVNDAGKMYTIDGAYKIFQMDHTPTINGSS  
ETPKYFVSROOKRTSGHITVSDHFKEMAKQGMIGNLYVALNAEGWSSGADVTL

terminator 4535..4872  
misc\_feature 4895  
/note="from rbcS gene"  
/note="SP6 RNA polymerase transcription initiation site"  
ORIGIN  
Query Match 14.5%; Score 171; DB 12; Length 4895;  
Best Local Similarity 100.0%; Pred No. 1.9e-87; Indels 0; Gaps 0;  
Matches 171; Conservative 0; Mismatches 0;  
QY 163 TTTTATATATGTGAAAAGAGTAATTTCACTTGGGCCACCTTTTATTACCGATATTTT 222  
| | | | |  
Db 2593 TTTTATATATGTGAAAAGAGTAATTTCACTTGGGCCACCTTTTATTACCGATATTTT 2534  
| | | | |  
QY 223 ACTTTATACCACTTTTAACTGATGTTTCACTTTTGACCAAGGTAATCTTACCTTTGTTT 282  
| | | | |  
Db 2533 ACTTTATACCACTTTTAACTGATGTTTCACTTTTGACCAAGGTAATCTTACCTTTGTTT 2474  
| | | | |  
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 333  
| | | | |  
Db 2473 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACTCGAGTA 2423  
| | | | |  
RESULT 11  
AX384394/c 9359 bp DNA linear PAT 19-MAR-2002  
LOCUS Sequence 3 from Patent WO214524.  
DEFINITION AX384394  
ACCESSION AX384394  
VERSION AX384394.1 GI:19577678  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS van der Valk,P., van Dun,C.M., Smeekens,S.C. and Proveniers,M.C.  
TITLE Inhibition of generative propagation in genetically modified  
herbicide resistant grasses  
JOURNAL Patent: WO 0214524-A 3 21-FEB-2002;  
Avantia Seeds B.V. (NL)  
FEATURES  
source  
1. .9359  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
839..1699  
CDS  
/note="unnamed protein product; Beta-lactamase gene (Ampr)"  
/codon\_start=1  
/translacion="MSIOHFRVALIPFPAAFCLEVFAPHPETLVKVKDAEDQLGARVGY  
IETDLSGKILIESFRPERFPMWSTFVLLCGAVLSRIDAQEQQLGRRIHYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAIIAMTSNDTANLLITIGDKELTAEFLHNMGDVTRL  
DRPEPELNEAIPIRDERPTTPVMATTLRKILTGSELITLSRQQLIDMEADKVAGPL  
LRSLAPGMFLPADSGAGERSRGIITAAIGDPKRSRIIVITTTGSAQTMDERRRQIA  
ETGASLIKHM"  
2941..4920  
promoter  
/note="ubi-promoter from maize"  
4921..6400  
misc\_feature  
/note="Ath gene from Arabidopsis thaliana"  
6401..6672  
polyA\_signal  
/note="Poly-A signal from the nopaline synthetase gene  
from Agrobacteriu m tumefaciens"  
7434..8084  
misc\_feature  
/note="first exon-intron combination from Ubi-maize"  
8085..9119  
misc\_feature  
/note="Hygromycin resistance gene from Escherichia coli"  
9120..9359  
ORIGIN  
Query Match 14.5%; Score 171; DB 6; Length 9359;





```
ORGANISM      synthetic construct
REFERENCE      other sequences; artificial sequences.
1
AUTHORS        Bakker,H.A., Florack,D.E. and Bosch,H.J.
TITLE          GntIIi (udp-n-acetylglucosamine:beta-d-mannoside
JOURNAL        beta(1,4)-n-acetylglucosaminyltransferase ii) expression in plants
                Patent: WO 03078614-A 10 25-SEP-2003;
                Plant Research International B.V. (NL)
FEATURES
source         1..7545
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Synthetic"

ORIGIN
Query Match      14.3%; Score 169; DB 6; Length 7545;
Best Local Similarity 100.0%; Pred. No. 3e-86;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      163 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACCGATATTT 222
        |||||||
DB      1473 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACCGATATTT 1414

QY      223 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 282
        |||||||
DB      1413 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 1354

QY      283 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 331
        |||||||
DB      1353 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 1305

RESULT 15
AX840289/c     11643 bp      DNA      linear      PAT 16-DEC-2003
LOCUS          AX840289
DEFINITION     Sequence 11 from Patent WO03078614.
ACCESSION      AX840289
VERSION        AX840289.1 GI:39978688
KEYWORDS
SOURCE         synthetic construct
ORGANISM       synthetic construct
OTHER SEQUENCES; artificial sequences.
1
REFERENCE      1
AUTHORS        Bakker,H.A., Florack,D.E. and Bosch,H.J.
TITLE          GntIIi (udp-n-acetylglucosamine:beta-d-mannoside
JOURNAL        beta(1,4)-n-acetylglucosaminyltransferase ii) expression in plants
                Patent: WO 03078614-A 11 25-SEP-2003;
                Plant Research International B.V. (NL)
FEATURES
source         1..11643
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Synthetic"

ORIGIN
Query Match      14.3%; Score 169; DB 6; Length 11643;
Best Local Similarity 100.0%; Pred. No. 3.2e-86;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      163 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACCGATATTT 222
        |||||||
DB      5573 TTTTATATAGTAGAAAAGATAATTTCACCTTGGGCCACCTTTATATACCGATATTT 5514

QY      223 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 282
        |||||||
DB      5513 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACAGGTAATCTTACCTTGT 5454

QY      283 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 331
        |||||||
DB      5453 TATTTTGACTATCCGACTCTCTTCACAGCATATGATGACCTCGAG 5405
```

Search completed: February 10, 2005, 13:13:42  
Job time : 7299.97 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 514.751 Seconds  
(without alignment)  
5727.097 Million cell updates/sec

Title: US-10-790-430-7

Perfect score: 498  
Sequence: 1 aatcgatcccaaatcgcgcac.....aaaaggtgcccacaagtcaa 498

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn1980s:\*\n2: geneseqn1990s:\*\n3: geneseqn2000s:\*\n4: geneseqn2001as:\*\n5: geneseqn2001bs:\*\n6: geneseqn2002as:\*\n7: geneseqn2002bs:\*\n8: geneseqn2003as:\*\n9: geneseqn2003bs:\*\n10: geneseqn2003cs:\*\n11: geneseqn2003ds:\*\n12: geneseqn2004as:\*\n13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 498   | 100.0       | 498    | 6  | ABK15244    |
| 2          | 118   | 23.7        | 1183   | 6  | ABK15245    |
| 3          | 101   | 20.3        | 2378   | 3  | AAD01016    |
| 4          | 70    | 14.1        | 1501   | 3  | AAC87194    |
| 5          | 70    | 14.1        | 1501   | 3  | AAC89325    |
| 6          | 70    | 14.1        | 2181   | 10 | ADC84543    |
| 7          | 70    | 14.1        | 2199   | 2  | AAT80052    |
| 8          | 70    | 14.1        | 5510   | 2  | AAQ12707    |
| 9          | 70    | 14.1        | 5643   | 2  | AAT80055    |
| 10         | 70    | 14.1        | 9359   | 6  | ABK1039     |
| 11         | 70    | 14.1        | 9359   | 6  | ABK10687    |
| 12         | 68    | 13.7        | 1259   | 6  | ADG69912    |
| 13         | 68    | 13.7        | 1259   | 11 | ADN88926    |
| 14         | 68    | 13.7        | 1392   | 2  | AAT80053    |
| 15         | 68    | 13.7        | 3657   | 6  | ADG69892    |
| 16         | 68    | 13.7        | 4604   | 6  | ADG69905    |
| 17         | 68    | 13.7        | 5118   | 12 | ADP73930    |
| 18         | 68    | 13.7        | 5164   | 11 | ADN88919    |
| 19         | 68    | 13.7        | 5267   | 11 | ADN88907    |
| 20         | 68    | 13.7        | 5277   | 6  | ADG69894    |

|    |    |      |       |    |          |                    |
|----|----|------|-------|----|----------|--------------------|
| 21 | 68 | 13.7 | 5277  | 11 | ADN88909 | Adn88909 pTBS-1 ve |
| 22 | 68 | 13.7 | 5295  | 6  | ADG69907 | Adg69907 Vector pu |
| 23 | 68 | 13.7 | 5295  | 11 | ADN88922 | Adn88922 pUG4 vect |
| 24 | 68 | 13.7 | 5327  | 6  | ADG69896 | Adg69896 Vector pt |
| 25 | 68 | 13.7 | 5327  | 11 | ADN88911 | Adn88911 pTP42 ve  |
| 26 | 68 | 13.7 | 5337  | 6  | ADG69888 | Adg69888 Vector pu |
| 27 | 68 | 13.7 | 5337  | 6  | ADG69900 | Adg69900 Vector pt |
| 28 | 68 | 13.7 | 5337  | 11 | ADN88914 | Adn88914 pTUS vect |
| 29 | 68 | 13.7 | 5337  | 11 | ADN88903 | Adn88903 pTUA vect |
| 30 | 68 | 13.7 | 5338  | 6  | ADG69884 | Adg69884 Vector pt |
| 31 | 68 | 13.7 | 5338  | 6  | ADG69898 | Adg69898 Vector pt |
| 32 | 68 | 13.7 | 5338  | 11 | ADN88913 | Adn88913 pTP3-1 ve |
| 33 | 68 | 13.7 | 5338  | 11 | ADN88899 | Adn88899 pTP10-1 v |
| 34 | 68 | 13.7 | 5345  | 6  | ADG69886 | Adg69886 Vector pu |
| 35 | 68 | 13.7 | 5345  | 11 | ADN88901 | Adn88901 pUA-4 ve  |
| 36 | 68 | 13.7 | 5387  | 6  | ADG69910 | Adg69910 Vector pt |
| 37 | 68 | 13.7 | 5387  | 11 | ADN88925 | Adn88925 pTP11-1 v |
| 38 | 68 | 13.7 | 5395  | 6  | ADG69890 | Adg69890 Vector pt |
| 39 | 68 | 13.7 | 5395  | 11 | ADN88905 | Adn88905 pTBS-14 v |
| 40 | 68 | 13.7 | 7545  | 10 | ACF58343 | Act58343 Nucleotid |
| 41 | 68 | 13.7 | 11643 | 10 | ACF58344 | Act58344 Nucleotid |
| 42 | 68 | 13.7 | 13680 | 12 | ADP73931 | Adp73931 Plasmid p |
| 43 | 67 | 13.5 | 4032  | 2  | AAV23237 | Aav23237 Plasmid p |
| 44 | 63 | 12.7 | 399   | 6  | AAV28024 | Aav28024 Chimera o |
| 45 | 62 | 12.4 | 6865  | 4  | AAF80287 | Aaf80287 Nucleotid |

## ALIGNMENTS

|          |   |                                |
|----------|---|--------------------------------|
| RESULT 1 | ABK15244  | standard; DNA; 498 BP.         |
| ID       | ABK15244;   |                                |
| AC       | ABK15244;   |                                |
| XX       | 08-MAY-2002   | (first entry)                  |
| DT       |   |                                |
| XX       | Corn nk603 5' transgene/genomic sequence PCR fragment.              |                                |
| DE       |   |                                |
| XX       | de; rice actin 1 promoter; RAI; RAI intron;                         |                                |
| KW       | chloroplast transit peptide gene; glyphosate resistance; corn;      |                                |
| KW       | 5-enol-pyruvylbiphosphate-3-phosphate synthase; EPPS; Hsp70 intron; |                                |
| KW       | transcriptional terminator; cauliflower mosaic virus 35S promoter;  |                                |
| KW       | PV-ZWGTJ2; transgenic; nk603.                                       |                                |
| XX       |   |                                |
| OS       | Zea mays.   |                                |
| OS       | Oryza sativa.   |                                |
| OS       | Synthetic.  |                                |
| OS       | Chimeric.   |                                |
| XX       |   |                                |
| FX       | Key   | Location/Qualifiers            |
| FT       | FT  | 1..304                         |
| FT       | FT  | /*tag= a                       |
| FT       | FT  | /note= "Corn genomic sequence" |
| FT       | FT  | 305..349                       |
| FT       | FT  | /*tag= b                       |
| FT       | FT  | /note= "Vector sequence"       |
| FT       | FT  | 350..498                       |
| FT       | FT  | /*tag= c                       |
| FT       | FT  | /note= "Rice ACT1 sequence"    |
| PN       | EP1167531-A1.   |                                |
| XX       |   |                                |
| PD       | 02-JAN-2002.  |                                |
| XX       |   |                                |
| PP       | 15-JUN-2001; 2001EP-00202314.                                       |                                |
| XX       |   |                                |
| PR       | 22-JUN-2000; 2000US-0213567P.                                       |                                |
| PR       | 13-OCT-2000; 2000US-0240014P.                                       |                                |
| PR       | 13-OCT-2000; 2000US-0241215P.                                       |                                |
| XX       |   |                                |
| PA       | (MONS ) MONSANTO TECHNOLOGY LLC.                                    |                                |





PT used to produce transgenic plants e.g. banana, wheat, maize or rice,  
 PR having resistance or tolerance to glyphosate herbicide.

PS Claim 19, Page 55-56; 87pp; English.

CC The invention relates to rice 5-enolpyruvylshikimate phosphate synthase  
 CC (EPSPS) genomic DNA (AAC87188). The invention also relates to an  
 CC expression cassette comprising, in the 5'-3' direction, one or more  
 CC transcriptional enhancer elements selected from AAC87190-C87196), the  
 CC rice EPSPS promoter, genomic DNA encoding a rice EPSPS chloroplast  
 CC transit peptide, genomic DNA encoding a EPSPS protein modified such that  
 CC it is resistant to glyphosate (AAC87189), and a transcriptional  
 CC terminator. The glyphosate resistant EPSPS contains a region (AAB29793)  
 CC containing two amino acid substitutions relative to the corresponding  
 CC wild-type region (AAB29792). The invention also encompasses plant genomic  
 CC EPSPS sequences identified via screening with a rice EPSPS intronic  
 CC sequence; vectors and host plant cells comprising a nucleic acid sequence  
 CC of the invention; transgenic plants (and tissues and seeds thereof)  
 CC comprising a nucleic acid sequence of the invention, optionally further  
 CC transformed with a DNA encoding an insect, fungal, viral, bacterial,  
 CC nematode, stress or herbicide resistance protein; and methods of  
 CC producing the transgenic plants of the invention. The nucleic acids and  
 CC constructs of the invention are used to produce a wide variety of  
 CC morphologically normal, glyphosate resistant plants. The glyphosate  
 CC resistant plants produced are particularly maize, soybean, cotton,  
 CC sugarcane and canola, but also other field crops, fruits and vegetables,  
 CC turf and forage grasses and nut-producing plants. The plants are  
 CC optionally resistant to insects, fungi, viruses, bacteria, nematodes,  
 CC stress, desiccation and/or other herbicides. They can be used in the  
 CC production of a herbicidal target for the high throughput in vitro  
 CC screening of potential herbicides. The present sequence represents an  
 CC enhancer element which may be used in the rice EPSPS expression cassette  
 CC of the invention

SQ Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;

Query Match 14.1%; Score 70; DB 3; Length 1501;

Best Local Similarity 100.0%; Pred. No. 1.8e-25;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 415

DB 750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 809

QY 416 TAAGATTACC 425

DB 810 TAAGATTACC 819

RESULT 5

AAC89325 standard; DNA; 1501 BP.

AAC89325;

07-MAR-2001 (first entry)

Rice actin enhancer.

5-enolpyruvylshikimate phosphate synthase; EPSPS; herbicide resistance;

glyphosate; ds.

Oryza sp.

MO200066747-A1.

09-NOV-2000.

20-APR-2000; 2000WO-GB0015172.

28-APR-1999; 99GB-00009967.

29-APR-1999; 99GB-00009969.

29-APR-1999; 99GB-00009972.

PR 29-APR-1999; 99GB-00009981.

PR 29-JUL-1999; 99GB-00017835.

PR 29-JUL-1999; 99GB-00017836.

PR 29-JUL-1999; 99GB-00017843.

PR 21-DEC-1999; 99GB-00030202.

PR 21-DEC-1999; 99GB-00030210.

PR 21-DEC-1999; 99GB-00030212.

PA (ZENE ) ZENECA LTD.

PI Hawkes TR, Warner SAJ, Andrews CJ, Bachoo S, Pickertill AP;

WPI; 2000-679764/66.

Example 20; Page 54; 98pp; English.

The present invention relates to an Oryza sp. 5-enolpyruvylshikimate

phosphate synthase (EPSPS) gene. Vectors containing the gene may be used

to produce plant tissues and fertile whole plants which are substantially

tolerant or substantially resistant to glyphosate herbicide and to

produce a herbicidal target which is used for high throughput in vitro

screening of potential herbicides

SQ Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;

Query Match 14.1%; Score 70; DB 3; Length 1501;

Best Local Similarity 100.0%; Pred. No. 1.8e-25;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 415

DB 750 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGATGTCCTCAAAATTAACAAAG 809

QY 416 TAAGATTACC 425

DB 810 TAAGATTACC 819

RESULT 6

ADC84543 standard; DNA; 2181 BP.

ADC84543;

01-JAN-2004 (first entry)

Rice actin promoter encoding sequence.

expression cassette; ds.

Melanomys caliginosus.

MO2003027257-A2.

03-APR-2003.

27-SEP-2002; 2002WO-US030895.

27-SEP-2001; 2001US-0325607P.

(SYGN ) SYNGENTA PARTICIPATIONS AG.

Levin JZ, Dietrich R, Budziszewski GJ;

WPI; 2003-354651/33.

New isolated nucleic acid molecule encoding a polypeptide with a 3'-5'

exonuclease domain, useful in molecular biology and transformation, in

particular for reproducing and predictably manipulating gene expression

PT in a plant cell.  
 XX  
 PS Disclosure; Page 106-107; 108bp; English.  
 XX  
 CC The present invention relates to an invention that results in the  
 CC expression of the endogenous nucleotide sequence in plant cell being  
 CC increased. In this method the plant cell comprises a second expression  
 CC cassette comprising a nucleic acid molecule, where the expression of the  
 CC nucleic acid molecule of interest in the plant cell is decreased as  
 CC compared to expression of the nucleic acid molecule of interest in a  
 CC plant cell lacking the first expression cassette. The methods and  
 CC compositions of the present invention are useful in molecular biology and  
 CC transformation, in particular for reproducing and predictably  
 CC manipulating gene expression in a plant cell. The present sequence is a  
 CC nucleotide sequence of rice actin promoter.  
 XX  
 SO Sequence 2181 BP; 602 A; 518 C; 430 G; 631 T; 0 U; 0 Other;  
 Query Match 14.1%; Score 70; DB 10; Length 2181;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-25;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 356 TACTCGAGTCATTCAATGCTTGAAGAAGAGTCGGGATAGTCCAAATPAAACAAGG 415  
 DB 755 TACTCGAGTCATTCAATGCTTGAAGAAGAGTCGGGATAGTCCAAATPAAACAAGG 814  
 QY 416 TAAGATTACC 425  
 DB 815 TAAGATTACC 824  
 RESULT 7  
 AAT80052  
 ID AAT80052 standard; cDNA; 2199 BP.  
 XX  
 AC AAT80052;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 04-NOV-1997 (first entry)  
 XX  
 DE Rice actin 1 gene promoter region.  
 XX  
 KW Promoter; rice; actin 1 gene; cytoplasm; cytoskeleton; extension growth;  
 KW cytoplasmic streaming; cell division; Rac1; maize; Adh1 promoter; ss.  
 XX  
 OS Oryza sativa.  
 XX  
 PN US5641876-A.  
 XX  
 PD 24-JUN-1997.  
 XX  
 PF 27-OCT-1993; 93US-00144602.  
 XX  
 PR 05-JAN-1990; 90US-00461490.  
 PR 18-SEP-1991; 91US-00762680.  
 PA (CORR ) CORNELL RES FOUND INC.  
 XX  
 FI Wu R, Mcelroy D;  
 XX  
 DR WPI; 1997-340996/31.  
 XX  
 PT Nucleic acid containing the promoter of the rice actin-1 gene - used to  
 PT direct efficient expression of foreign genes in rice.  
 XX  
 PS Claim 3; Col 33-36; 29pp; English.  
 XX  
 CC This sequence represents the promoter region from the rice actin-1 (Rac1)  
 CC gene. Cytoplasmic actin is a fundamental and essential component of the  
 CC eukaryotic cell and cytoskeleton. In higher plant cells, a number of  
 CC cellular processes, such as cytoplasmic streaming, extension growth and  
 CC cell division are thought to involve the cytoskeletal actin protein. All  
 CC of the studied plant actins consist of four exons of conserved length,

CC separated by 3 introns of variable length. In rice there are at least 8  
 CC actin-like sequences per haploid genome. Rac1 encodes a transcript that  
 CC is relatively abundant in all rice tissues. This sequence is an example  
 CC of a nucleic acid molecule of the invention. The nucleic acid molecules  
 CC of the invention contain the Rac1 gene (or fragments) with promoter  
 CC activity in monocotyledonous plants. The promoter is used to direct  
 CC expression of foreign genes in transgenic rice and other plants. The  
 CC actin promoter is more efficient in rice transformation than previously  
 CC proposed promoters (e.g. 5 times more active than the maize Adh1  
 CC promoter) and has constitutive activity in space and time. (Updated on 25  
 CC -MAR-2003 to correct PF field.)  
 XX  
 SO Sequence 2199 BP; 603 A; 515 C; 446 G; 635 T; 0 U; 0 Other;  
 Query Match 14.1%; Score 70; DB 2; Length 2199;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-25;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 356 TACTCGAGTCATTCAATGCTTGAAGAAGAGTCGGGATAGTCCAAATPAAACAAGG 415  
 DB 806 TACTCGAGTCATTCAATGCTTGAAGAAGAGTCGGGATAGTCCAAATPAAACAAGG 865  
 QY 416 TAAGATTACC 425  
 DB 866 TAAGATTACC 875  
 RESULT 8  
 AAQ12707  
 ID AAQ12707 standard; DNA; 5510 BP.  
 XX  
 AC AAQ12707;  
 XX  
 DT 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 27-SEP-1991 (first entry)  
 XX  
 DE Rice actin gene Rac1.  
 XX  
 KW Promoter; transgenic plant; monocotyledonous; ss.  
 XX  
 OS Oryza sativa; variety IR26.  
 XX  
 FH Key  
 FT Location/Qualifiers  
 FT 1..2071  
 FT /\*tag= a  
 FT 1027..1038  
 FT /\*tag= b  
 FT 1078..1088  
 FT /\*tag= c  
 FT /\*tag= direct  
 FT 1301..1333  
 FT /\*tag= d  
 FT /\*tag= tandem (imperfect)  
 FT 1352..1386  
 FT /\*tag= e  
 FT /\*tag= tandem (imperfect)  
 FT /\*note= "eight pentamers"  
 FT 1465..1505  
 FT /\*tag= f  
 FT /\*tag= purine-rich region  
 FT 1609..1617  
 FT /\*tag= g  
 FT 1647..1653  
 FT /\*tag= h  
 FT /\*tag= transcription initiation signal  
 FT 1650..3741  
 FT /\*tag= p  
 FT /\*tag= genomic clone pRac1  
 FT 1650..1728  
 FT /\*tag= i  
 FT /\*note= "transcription initiation site"

```

FT repeat_region 1650. .1676
FT /*tag= j
FT /rpt_type= tandem
FT /note= "GC-rich - repeated (A/T)CC triplets"
FT misc_RNA 1728. .1731
FT /*tag= k
FT /label= 5' acceptor splice site
FT intron 1729. .2043
FT /*tag= l
FT misc_RNA 2008. .2012
FT /*tag= m
FT /label= putative branch point splice site
FT misc_RNA 2034. .2044
FT /*tag= n
FT /label= 3' donor splice site
FT CDS 2051. .3600
FT /*tag= o

XX PN WO9109948-A.
XX XX
XX PD 11-JUL-1991.
XX XX
XX PF 05-JAN-1990; 90US-00461490.
XX XX
XX PR 05-JAN-1990; 90US-00461490.
XX XX
XX PA (CORR ) CORNELL RES FOUND INC.
XX XX
XX PI Wu R, Mcelroy D;
XX DR WPI; 1991-222901/30.
XX XX
XX PT Rice actin gene promoter - used to confer high level expression of
XX foreign genes in transgenic rice and other plants.
XX XX
XX PS Claim 8; Page 10; 47pp; English.
XX XX
XX CC The seugence was obtd. from a clone isolated from a rice genomic library
XX screened with a heterologous actin probe. The CDS encodes a 377 amino
XX acid actin protein of 41.9 kD. The gene is interrupted by three introns
XX (position not disclosed). Analysis also revealed an additional intron 5'
XX of the translation initiation region. This 5' intron separates a 79 bp GC
XX rich 5' noncoding exon from the exon coding the translation initiation
XX codon. This is one of few cases of a 5' noncoding exon in a plant gene.
XX CC The promoter region can be used to confer high level expression of
XX foreign genes in transgenic rice and other agronomically important
XX CC plants. It is five times more active than the maize Adh1 promoter in
XX trans-formed rice. (Updated on 25-MAR-2003 to correct PA field.)
XX CC (Updated on 24-OCT-2003 to standardise OS field)
XX CC
XX SQ Sequence 5510 BP; 1466 A; 1262 C; 1163 G; 1619 T; 0 U; 0 Other;
XX XX

Query Match 14.1%; Score 70; DB 2; Length 5510;
Best Local Similarity 100.0%; Pred. No. 1.7e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCAATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATATAAACAAGG 415
DB 809 TACTCGAGTCATTCAATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATATAAACAAGG 868

QY 416 TAAGATTACC 425
DB 869 TAAGATTACC 878

RESULT 9
AAT80055
ID AAT80055 standard; cDNA; 5643 BP.
XX AC
XX AT80055;
XX XX
XX DT 25-MAR-2003 (revised)
DT 04-NOV-1997 (first entry)

```

```

XX DE Rice actin 1 gene.
XX DE
XX KM Promoter; rice; actin 1 gene; cytoplasm; cytoskeleton; extension growth;
XX cytoplasmic streaming; cell division; RAC1; maize; Adh1 promoter; ss.
XX OS
XX Oryza. sativa.
XX PN
XX US5641876-A.
XX PD 24-JUN-1997.
XX PF 27-OCT-1993; 93US-00144602.
XX XX
XX PR 05-JAN-1990; 90US-00461490.
XX PR 18-SEP-1991; 91US-00762680.
XX XX
XX PA (CORR ) CORNELL RES FOUND INC.
XX XX
XX PI Wu R, Mcelroy D;
XX DR WPI; 1997-340996/31.
XX XX
XX PT Nucleic acid containing the promoter of the rice actin-1 gene - used to
XX direct efficient expression of foreign genes in rice.
XX PS Example 1; Col 3-7; 29pp; English.
XX XX
XX CC This sequence represents a the rice actin-1 (Rac1) gene amplified using
XX AAT80054. Cytoplasmic actin is a fundamental and essential component of
XX the eukaryotic cell and cytoskeleton. In higher plant cells, a number of
XX cellular processes, such as cytoplasmic streaming, extension growth and
XX cell division are thought to involve the cytoskeletal actin protein. All
XX of the studied plant actins consist of four exons of conserved length,
XX separated by 3 introns of variable length. In rice there are at least 8
XX actin-like sequences per haploid genome. Rac1 encodes a transcript that
XX is relatively abundant in all rice tissues. This sequence is an example
XX of a nucleic acid molecule of the invention. The nucleic acid molecules
XX of the invention contain the Rac1 gene, or fragments with promoter
XX activity (see AAT80052 and AAT80053) in monocotyledonous plants. The
XX promoter is used to direct expression of foreign genes in transgenic rice
XX and other plants. The actin promoter is more efficient in rice
XX transformation than previously proposed promoters (e.g. 5 times more
XX active than the maize Adh1 promoter) and has constitutive activity in
XX space and time. (Updated on 25-MAR-2003 to correct PF field.)
XX CC
XX SQ Sequence 5643 BP; 1500 A; 1283 C; 1200 G; 1660 T; 0 U; 0 Other;
XX XX

Query Match 14.1%; Score 70; DB 2; Length 5643;
Best Local Similarity 100.0%; Pred. No. 1.7e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCAATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATATAAACAAGG 415
DB 809 TACTCGAGTCATTCAATATGCTTGAGAGAGAGTCGGAGTAGTCCAAATATAAACAAGG 868

QY 416 TAAGATTACC 425
DB 869 TAAGATTACC 878

RESULT 10
ABK11039
ID ABK11039 standard; DNA; 9359 BP.
XX AC
XX ABK11039;
XX XX
XX DT 05-JUN-2002 (first entry)
XX DE pVDH636 vector containing Arabidopsis Ath1 gene.
XX KM pVDH636 vector; circular; cyclic; grass; plant; herbicide resistance;
XX baseball; cricket; football; golf; rugby; soccer; tennis; lawn; park;

```

```
KW athletic field; animal feedstuff; grass flowering inhibitor; Ath1; gene;
XX da.
OS Arabidopsis thaliana.
OS Zea mays.
OS Agrobacterium tumefaciens.
OS Oryza sativa.
OS Escherichia coli.
OS Cauliflower mosaic virus.
OS Synthetic.
FH Key Location/Qualifiers
FT CDS 4930..6360
FT /*tag= a
FT /product= "Ath1"
FT /partial
FT /note= "No stop codon"
PN WO200214524-A2.
XX
XX
XX 21-FEB-2002.
PD
PF 16-AUG-2001; 2001WO-EP009572.
XX
XX 18-AUG-2000; 2000US-0226422P.
XX 27-NOV-2000; 2000US-0253274P.
XX 27-NOV-2000; 2000US-0253327P.
XX 22-JUN-2001; 2001US-0300220P.
XX
XX (ADVA-) ADVANTA SEEDS BV.
XX
XX Van Der Valk P, Van Dun CMP, Smeekens SCM, Proveniers MCG;
XX WPI: 2002-257611/30.
XX P-PSDB; AAU76514.
XX
XX New genetically modified grasses that express inhibited generative
XX propagation, or herbicide resistance, useful for forage (e.g. cattle
XX feedstuff) or amenity purposes (e.g. for use in an athletic field, lawn
XX or park).
XX
XX Example 1; Fig 6; 56pp; English.
XX
XX The invention relates to a grass plant, which has been genetically
XX modified to substantially inhibit generative propagation and carry
XX herbicide resistance. The grass is useful for growth and/or propagation
XX of grasses. The grass is used to play at least one sport (e.g. baseball,
XX cricket, football, golf, rugby, soccer or tennis), or used at least in a
XX portion of an athletic field, lawn or park. The grass is fed to animal
XX (e.g. cattle, goat, horse or sheep) or used as an animal feedstuff. The
XX present sequence represents the coding sequence of pVDH636 vector
XX containing the Arabidopsis Ath1 gene, which was used to inhibit flowering
XX in grasses
XX
XX Sequence 9359 BP; 2377 A; 2222 C; 2133 G; 2624 T; 0 U; 3 Other;
SQ
XX
XX Query Match 14.1%; Score 70; DB 6; Length 9359;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-25;
XX Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 356 TACTGAGGTCAATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGG 415
DB 6688 TACTGAGGTCAATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGG 6747
QY 416 TAAAGTTACC 425
DB 6748 TAAAGTTACC 6757
XX
XX RESULT 11
XX ABK10687 standard; DNA; 9359 BP.
XX
```

```
AC ABK10687;
XX
XX 05-JUN-2002 (first entry)
XX
XX Transformation vector plasmid pVDH636 DNA.
XX
XX Grass; plant; transgenic; flowering inhibition; inflorescence; gene; ds;
XX tiller production; delayed heading; gibberellic acid; phytohormone; golf;
XX genetically modified grass; athletic field; sport; baseball; cricket;
XX football; rugby; soccer; tennis; lawn; landscaping; cattle; horse; sheep;
XX goat; animal feedstuff; Arabidopsis thaliana; Ath1; homeobox gene;
XX circular; cyclic; pVDH636.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 4930..6351
XX /*tag= a
XX /product= "Arabidopsis thaliana Ath1"
XX
XX WO200214486-A2.
XX
XX
XX 21-FEB-2002.
PD
PF 16-AUG-2001; 2001WO-EP009570.
XX
XX 18-AUG-2000; 2000US-0226422P.
XX 27-NOV-2000; 2000US-0253274P.
XX 27-NOV-2000; 2000US-0253327P.
XX 22-JUN-2001; 2001US-0300220P.
XX
XX (ADVA-) ADVANTA SEEDS BV.
XX
XX Van Der Valk P, Van Dun CMP, Smeekens SCM, Proveniers MCG;
XX WPI: 2002-257603/30.
XX P-PSDB; AAU76885.
XX
XX New genetically modified grass useful for growing and/or propagating
XX grass in athletic fields (for sports such as baseball, cricket, football,
XX golf, rugby, soccer and tennis), or as animal feedstuff for cattle, goat,
XX horse and sheep.
XX
XX Example 1; Fig 6; 45pp; English.
XX
XX The invention relates to a grass plant which has been genetically
XX modified to substantially inhibit generative propagation. The genetic
XX modification may result in a heritable change in one or more plant
XX characteristics such as inhibition of flowering (or substantial delay
XX that amounts to inhibition), absence of inflorescence, increased
XX production of tillers, delayed heading and inhibition of the
XX developmental switch from vegetative to generative growth. A method of
XX making a grass involves transforming the grass with a nucleic acid which
XX interferes with metabolism of gibberellic acid. A grass can be treated by
XX applying a phytohormone to at least partially relieve or reverse a change
XX in plant characteristic resulting from genetic modification. The
XX genetically modified grass is useful for growing and/or propagating grass
XX in athletic fields (for sports such as baseball, cricket, football, golf,
XX rugby, soccer and tennis), lawns, parks and other types of landscaping.
XX The grass is also useful as an animal feedstuff for cattle, goats, horses
XX and sheep, due to its increased vegetative growth, improved digestibility
XX and/or nutritional value as animal feedstuff. This sequence represents
XX production of transgenic grasses expressing the Arabidopsis thaliana
XX homeobox gene, Ath1
XX
XX Sequence 9359 BP; 2377 A; 2222 C; 2133 G; 2624 T; 0 U; 3 Other;
SQ
XX
XX Query Match 14.1%; Score 70; DB 6; Length 9359;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-25;
XX Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 356 TACTGAGGTCAATTCATATGCTTGAGAGAGAGTCGGATAGTCCAAATAAACAAGG 415
XX
```

DB 6688 TACTCGAGTCATTCAATATGCTTGAAGAGAGTCGGATAGTCCAAATAAACAAGG 6747  
QY 416 TAAGATTACC 425  
DB 6748 TAAGATTACC 6757

RESULT 12  
ADG69912  
ID ADG69912 standard; DNA; 1259 BP.

AC ADG69912;

DT 11-MAR-2004 (first entry)

DE Actin promoter and nucleotide sequence.

transgenic plant; ferulic acid esterase; FAE; enzyme; phenolic acid;  
plant; cell wall; improved digestibility; biomass conversion;  
highly fermentable carbohydrate; gene; ds.

OS Synthetic.

PN WO200268666-A1.

PD 06-SEP-2002.

PP 16-NOV-2001; 2001WO-US043588.

PR 17-NOV-2000; 2000US-0249608P.

XX (GEMV ) GENENCOR INT INC.

PI Dunn-Coleman N, Langdon T, Morris P;

DR WPI; 2002-698675/75.

PT New transgenic plant comprising an expression cassette with a promoter  
operably linked to a ferulic acid esterase encoding polynucleotide;  
PT useful in improving digestibility for livestock and enhancing biomass  
conversion.

PS Disclosure; Fig 50A-B; 208pp; English.

XX The present invention describes a transgenic plant (i) comprising an  
expression cassette with a promoter operably linked to a ferulic acid  
esterase encoding polynucleotide. Also described: (1) controlling (M1)  
the level of phenolic acids in plant cell walls of a transgenic plant by  
introducing into the plant an expression cassette comprising a promoter  
operably linked to a ferulic acid esterase encoding polynucleotide; and  
(2) a transgenic plant (ii) produced by (M1). The transgenic plants are  
useful in improving digestibility for livestock and enhancing biomass  
conversion. The method is useful for enhancing the production of more  
highly fermentable carbohydrates in plants, especially forage grasses.  
The expression cassette is useful for controlling the level of phenolic  
acids in plant cell walls of a transgenic plant. The present sequence is  
used in the exemplification of the present invention.

XX Sequence 1259 BP; 337 A; 332 C; 268 G; 322 T; 0 U; 0 Other;

Query Match 13.7%; Score 68; DB 6; Length 1259;  
Best Local Similarity 100.0%; Pred. No. 2e-24;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCTATGCTTGAAGAGAGTCGGGATAGTCCAAATAAACAAGGTA 417  
DB 16 CTCGAGTCATTCTATGCTTGAAGAGAGTCGGGATAGTCCAAATAAACAAGGTA 75

QY 418 AGATTACC 425  
DB 76 AGATTACC 83

RESULT 13

ID ADN88926  
ADN88926 standard; DNA; 1259 BP.

AC ADN88926;

DT 01-JUL-2004 (first entry)

DE Actin promoter.

transgenic plant; ferulic acid esterase; fermentable carbohydrate;  
animal feed; phenolic acid; actin; promoter; ds.

OS Synthetic.

PN WO2003043411-A2.

PD 30-MAY-2003.

PP 21-MAY-2002; 2002WO-US016239.

PR 16-NOV-2001; 2001US-00991209.

PR 16-NOV-2001; 2001WO-US043588.

XX (GEMV ) GENENCOR INT INC.

PI Dunn-Coleman N, Langdon T, Morris P;

DR WPI; 2003-457563/43.

PT New transgenic plant comprising an expression cassette with a promoter  
operably linked to ferulic acid esterase encoding polynucleotide; useful  
for enhancing availability of fermentable carbohydrates as feed for grass  
-fed animals.

PS Disclosure; Fig 50; 190pp; English.

XX The present invention relates to transgenic plants comprising an  
expression cassette having a promoter operably linked to a ferulic acid  
esterase (FAE) encoding polynucleotide. The transgenic plants are useful  
for enhancing the availability of fermentable carbohydrates as feed for  
grass-fed animals. The methods are useful for reducing the levels of  
phenolic acids in the cell walls available for cross-linking either by  
directly disrupting ester bonds linking phenolics and lignins to cell  
wall polysaccharides or by preventing excessive feruloylation of cell wall  
carbohydrates. The present sequence was used to illustrate the invention.

XX Sequence 1259 BP; 337 A; 332 C; 268 G; 322 T; 0 U; 0 Other;

Query Match 13.7%; Score 68; DB 11; Length 1259;  
Best Local Similarity 100.0%; Pred. No. 2e-24;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCTATGCTTGAAGAGAGTCGGGATAGTCCAAATAAACAAGGTA 417  
DB 16 CTCGAGTCATTCTATGCTTGAAGAGAGTCGGGATAGTCCAAATAAACAAGGTA 75

QY 418 AGATTACC 425  
DB 76 AGATTACC 83

RESULT 14

ID AAT80053  
AAT80053 standard; cDNA; 1392 BP.

AC AAT80053;

DT 25-MAR-2003 (revised)

DT 04-NOV-1997 (first entry)

DE Rice actin 1 gene promoter fragment.



XX Promoter; rice; actin 1 gene; cytoplasm; cytoskeleton; extension growth;  
 KW cytoplasmic streaming; cell division; Rac1; maize; Adh1 promoter; ss.  
 XX  
 OS *Oryza sativa*.  
 XX  
 PN US561876-A.  
 XX  
 PD 24-JUN-1997.  
 XX  
 PF 27-OCT-1993; 93US-00144602.  
 XX  
 PR 05-JAN-1990; 90US-00461490.  
 PR 18-SEP-1991; 91US-00762680.  
 PA (CORR ) CORNELL RES FOUND INC.  
 XX  
 PI Wu R, McElroy D;  
 DR WPI; 1997-340996/31.  
 XX  
 PT Nucleic acid containing the promoter of the rice actin-1 gene - used to  
 PT direct efficient expression of foreign genes in rice.  
 XX  
 PS Claim 4; Col 35-36; 29pp; English.  
 XX  
 SS This sequence represents a fragment from the promoter region of the rice  
 CC actin-1 (Rac1) gene. Cytoplasmic actin is a fundamental and essential  
 CC component of the eukaryotic cell and cytoskeleton. In higher plant cells,  
 CC a number of cellular processes, such as cytoplasmic streaming, extension  
 CC growth and cell division are thought to involve the cytoskeletal actin  
 CC protein. All of the studied plant actins consist of four exons of  
 CC conserved length, separated by 3 introns of variable length. In rice  
 CC there are at least 8 actin-like sequences per haploid genome. Rac1  
 CC encodes a transcript that is relatively abundant in all rice tissues.  
 CC This sequence is an example of a nucleic acid molecule of the invention.  
 CC The nucleic acid molecule of the invention contains the Rac1 gene (or  
 CC fragments) with promoter activity in monocotyledonous plants. The  
 CC promoter is used to direct expression of foreign genes in transgenic rice  
 CC and other plants. The actin promoter is more efficient in rice  
 CC transformation than previously proposed promoters (e.g. 5 times more  
 CC active than the maize Adh1 promoter) and has constitutive activity in  
 CC space and time. (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 SQ Sequence 1392 BP; 376 A; 343 C; 309 G; 364 T; 0 U; 0 Other;  
 XX  
 QY Query Match 13.7%; Score 68; DB 2; Length 1392;  
 Db Best Local Similarity 100.0%; Pred. No. 1.9e-24;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 358 CTCGAGGTCAATTCATATGCTTGAGAGAGATCGGATAGTCCAAATTAACAAGGTA 417  
 Db 1 CTCGAGGTCAATTCATATGCTTGAGAGAGATCGGATAGTCCAAATTAACAAGGTA 60  
 QY 418 AGATTACC 425  
 Db 61 AGATTACC 68  
 XX  
 RESULT 15  
 ADG69892  
 ID ADG69892 standard; DNA; 3657 BP.  
 XX  
 AC ADG69892;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Vector pTP8-5 nucleotide sequence.  
 XX  
 KM transgenic plant; ferulic acid esterase; PAB; enzyme; phenolic acid;  
 KM plant; cell wall; improved digestibility; biomass conversion;  
 KM highly fermentable carbohydrate; gene; ds.  
 XX

OS Synthetic.  
 XX  
 PN WO200268666-A1.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 16-NOV-2001; 2001WO-US043588.  
 XX  
 PR 17-NOV-2000; 2000US-0249608P.  
 PA (GENENCOR ) GENENCOR INT INC.  
 XX  
 PI Dunn-Coleman N, Langdon T, Morris P;  
 DR WPI; 2002-698675/75.  
 DR P-PSDB; ADG69893.  
 XX  
 PT New transgenic plant comprising an expression cassette with a promoter  
 PT operably linked to a ferulic acid esterase encoding polynucleotide,  
 PT useful in improving digestibility for livestock and enhancing biomass  
 PT conversion.  
 XX  
 PS Disclosure; Fig 36B-E; 208pp; English.  
 XX  
 SS The present invention describes a transgenic plant (I) comprising an  
 CC expression cassette with a promoter operably linked to a ferulic acid  
 CC esterase encoding polynucleotide. Also described: (1) controlling (M1)  
 CC the level of phenolic acids in plant cell walls of a transgenic plant by  
 CC interducing into the plant an expression cassette comprising a promoter  
 CC operably linked to a ferulic acid esterase encoding polynucleotide; and  
 CC (2) a transgenic plant (II) produced by (M1). The transgenic plants are  
 CC useful in improving digestibility for livestock and enhancing biomass  
 CC conversion. The method is useful for enhancing the production of more  
 CC highly fermentable carbohydrates in plants, especially forage grasses.  
 CC The expression cassette is useful for controlling the level of phenolic  
 CC acids in plant cell walls of a transgenic plant. The present sequence is  
 CC used in the exemplification of the present invention.  
 CC  
 SQ Sequence 3657 BP; 916 A; 1012 C; 868 G; 860 T; 0 U; 1 Other;  
 XX  
 QY Query Match 13.7%; Score 68; DB 6; Length 3657;  
 Db Best Local Similarity 100.0%; Pred. No. 1.9e-24;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 358 CTCGAGGTCAATTCATATGCTTGAGAGAGATCGGATAGTCCAAATTAACAAGGTA 417  
 Db 2430 CTCGAGGTCAATTCATATGCTTGAGAGAGATCGGATAGTCCAAATTAACAAGGTA 2489  
 QY 418 AGATTACC 425  
 Db 2490 AGATTACC 2497  
 Search completed: February 10, 2005, 13:43:58  
 Job time : 516.751 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd

QM nucleic - nucleic search, using sw model

```
Run on:      February 10, 2005, 07:47:34 ; Search time 7997.05 Seconds
              (without alignments)
              5630.828 Million cell updates/sec
```

Title: US-10-790-430-8

Sequence: 1 gacgttattcatgagatggg.....aaggaaggcgcgagatgatg 1183

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

Database

```

1:  gb_est1:*
2:  gb_est2:*
3:  gb_hc:*
4:  gb_est3:*
5:  gb_est4:*
6:  gb_est5:*
7:  gb_est6:*
8:  gb_gss1:*
9:  gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID        | Description          |
|------------|-------|-------------|--------|----|-----------|----------------------|
| C 1        | 305   | 25.8        | 537    | 8  | B2310383  | B2310383 i5c35c07.b  |
| C 2        | 305   | 25.8        | 553    | 8  | BH883307  | BH883307 hW43f05.g   |
| C 3        | 305   | 25.8        | 648    | 8  | B2314790  | B2314790 h2z2a01.b   |
| C 4        | 305   | 25.8        | 680    | 8  | B2625170  | B2625170 i996e05.g   |
| C 5        | 305   | 25.8        | 717    | 8  | B2320382  | B2320382 h07h12.g    |
| C 6        | 305   | 25.8        | 738    | 8  | CC161210  | CC161210 i176c10.b   |
| C 7        | 305   | 25.8        | 776    | 8  | BH878576  | BH878576 h82h12.b    |
| C 8        | 305   | 25.8        | 824    | 6  | CD438581  | CD438581 EL01N0515   |
| C 9        | 305   | 25.8        | 920    | 8  | CC357906  | CC357906 P0H0367ED   |
| C 10       | 280   | 23.7        | 600    | 8  | B2375833  | B2375833 i62p08.g    |
| C 11       | 275   | 23.2        | 600    | 8  | B2373363  | B2373363 i67f08.b    |
| C 12       | 273   | 23.1        | 313    | 8  | CC025812  | CC025812 359l.1.46g  |
| C 13       | 256   | 21.6        | 646    | 8  | B2359336  | B2359336 i72d02.b    |
| C 14       | 253   | 21.4        | 650    | 8  | B2372017  | B2372017 i62p08.b    |
| C 15       | 241   | 20.4        | 428    | 8  | B2637555  | B2637555 i9D0d01.g   |
| C 16       | 238   | 20.1        | 835    | 8  | B2697563  | B2697563 PUCCK44TfD  |
| C 17       | 238   | 20.1        | 955    | 8  | B2747937  | B2747937 PUCS09TfT   |
| C 18       | 195   | 16.5        | 1010   | 9  | CM002222  | CM002222 ZMBBHG0000  |
| C 19       | 184   | 15.6        | 736    | 9  | CC804949  | CC804949 i1h8d04.g   |
| C 20       | 183   | 15.5        | 453    | 8  | B2310528  | B2310528 i65D09.g    |
| C 21       | 183   | 15.5        | 606    | 8  | B2778381  | B2778381 i6c3p07.g   |
| C 22       | 182   | 15.4        | 613    | 8  | CC156828  | CC156828 i615h05.g   |
| C 23       | 177   | 15.0        | 1047   | 9  | CM001399  | CM001399 ZMBBHG0000  |
| C 24       | 177   | 15.0        | 1060   | 9  | CL1990345 | CL1990345 ZMBBHF0000 |

|   |    |      |      |      |   |          |          |             |
|---|----|------|------|------|---|----------|----------|-------------|
| C | 25 | 1.17 | 15.0 | 1062 | 9 | CM002030 | CM002030 | ZMMBH000    |
| C | 26 | 1.19 | 14.3 | 1013 | 6 | CM002257 | ZMMBH000 |             |
| C | 27 | 1.68 | 14.2 | 670  | 6 | CB867772 | CB867772 | HC01D06w    |
| C | 28 | 1.66 | 14.0 | 166  | 8 | BH415252 | BH415252 | 1007042C0   |
| C | 29 | 1.60 | 13.5 | 672  | 8 | BZ349476 | BZ349476 | h-r41b05-g  |
| C | 30 | 1.60 | 13.5 | 955  | 9 | CL239469 | CL123946 | h-r41b05-g  |
| C | 31 | 1.59 | 13.4 | 500  | 8 | BZ316096 | BZ316096 | 1a67f08.b   |
| C | 32 | 1.59 | 13.4 | 500  | 8 | BZ322609 | BZ322609 | 1a67f08.g   |
| C | 33 | 1.58 | 13.4 | 500  | 8 | CM001948 | CM001948 | ZMMBH000    |
| C | 34 | 1.52 | 12.8 | 633  | 9 | CG159311 | CG159311 | PUIXK31TD   |
| C | 35 | 1.45 | 12.3 | 568  | 8 | BZ341607 | BZ341607 | 1c46f03.g   |
| C | 36 | 1.45 | 12.3 | 606  | 8 | BZ329855 | BZ329855 | hvr8d04.g   |
| C | 37 | 1.45 | 12.3 | 632  | 8 | BZ332275 | BZ332275 | h-r266e12.g |
| C | 38 | 1.45 | 12.3 | 707  | 9 | CM006474 | CM006474 | ZMMBL000    |
| C | 39 | 1.45 | 12.3 | 782  | 8 | BZ781027 | BZ781027 | 1j32b10.g   |
| C | 40 | 1.45 | 12.3 | 845  | 9 | CM009185 | CM009185 | ZMMBL001    |
| C | 41 | 1.34 | 11.3 | 474  | 7 | CM061066 | CM061066 | BNEL6468    |
| C | 42 | 1.34 | 11.3 | 480  | 7 | CV059547 | CV059547 | BNEL49C6    |
| C | 43 | 1.34 | 11.3 | 529  | 7 | CV057962 | CV057962 | BNEL3258    |
| C | 44 | 1.34 | 11.3 | 544  | 7 | CV060284 | CV060284 | BNEL56d9    |
| C | 45 | 1.34 | 11.3 | 549  | 7 | CV060835 | CV060835 | BNEL61g2    |

## ALIGNMENTS

| RESULT 1                  | LOCUS                                      | DEFINITION      |
|---------------------------|--|-----------------|
| B2310383                  | 537 bp                                     | DNA linear      |
| B2310383                  |  | GSS 06-NOV-2003 |
| IC53C07.D1                | WGS-ZmayBF (JM107 adapted methyl) filtered | Zee may's       |
| genomic clone IC53C07 5', | genomic survey sequence.                   |                 |

| SOURCE ORGANISM | Zea mays | Zea mays |
|-----------------|----------|----------|
|                 |          |          |

REFERENCE  
AUTHORS  
1 (bases 1 to 537)  
Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,  
Wasserman, I.

**TITLE** Genomic shotgun sequences from Zea mays (metyl-1-lilleieu)  
**JOURNAL** unpublished (2002)  
**COMMENT** Contact: W. Richard McCombie  
Genotyping Center

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. .537             |

/clone.lib="MGS-zmavf (JM107 adapted methyl filtered)  
 /note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I  
 The vector was digested with XbaI and one nucleotide was  
 added by fill in in the recessive 3' end. The genomic DNA  
 was nebulized, end repaired, adaptor ligated and size  
 fractionated using sephadex. The resulting fragments were  
 between 0.8 and 3 kb and were cloned into the vector  
 (x/y reads in M13mp19, b/g reads in pUC19). The same  
 ligation was transformed either JM107 or DH5a. "

## ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 537;

Best Local Similarity 100.0%; Pred. No. 2.5e-157; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5a. "

382 ACAACCTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 441  
DB 91 ACAACCTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 150  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

442 TAGATATTATGATCATTTGATGCTTTATTTCTCTGAAAGCGTTTCATTTTTTTTAC 501  
DB 151 TAGATATTATGATCATTTGATGCTTTATTTCTCTGAAAGCGTTTCATTTTTTTTAC 210

502 AGACGCTTTTATAGAGGTGACATCATTAATGCGCATAGGTGTATACATCGGTATA 561  
DB 211 AGACGCTTTTATAGAGGTGACATCATTAATGCGCATAGGTGTATACATCGGTATA 270

562 CAACCTTAACCGTACACCACTTTAGCAATGAGCTGTAAATGCGCATCTCTTCGCTACCA 621  
DB 271 CAACCTTAACCGTACACCACTTTAGCAATGAGCTGTAAATGCGCATCTCTTCGCTACCA 330

622 GCACCTTTTACCAATACCTTCGCTGCTTGCAGAAACCACTGTACCAATAGCATCTACT 681  
DB 331 GCACCTTTTACCAATACCTTCGCTGCTTGCAGAAACCACTGTACCAATAGCATCTACT 390

682 GTTCT 686  
DB 391 GTTCT 395

RESULT 2  
BH883307/c 553 bp DNA linear GSS 05-AUG-2002  
LOCUS h43f05.g1 WGS-Zmaysef (JM107 adapted methyl filtered) Zea mays  
DEFINITION genomic clone h43f05 5', genomic survey sequence.

ACCESSION BH883307  
VERSION BH883307.1 GI:22119204  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 553)  
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,  
Katzemburger, F., King, L., Miller, B., Muller, S., Nascento, L.,  
Zutavern, T., McCombie, W.R., and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)

TITLE Lita Annenberg Hazen Genome Sequencing Center  
JOURNAL Contact: W. Richard McCombie  
COMMENT Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: h43 row: f column: 05  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 553.  
Location/Qualifiers  
1. .553  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
/clone="h43f05"  
/lab host="JM107 or DH5a"  
/note="Organ: immature ears; site\_1: Xba I; site\_2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA

FEATURES  
source

382 ACAACCTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 441  
DB 442 ACAACCTTCCACTCTAGTGTGAGTGATCCTGTTATCTCTTCGAACCATTAACAGAC 383  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

442 TAGATATTATGATCATTTGATGCTTTATTTCTCTGAAAGCGTTTCATTTTTTTTAC 501  
DB 382 TAGATATTATGATCATTTGATGCTTTATTTCTCTGAAAGCGTTTCATTTTTTTTAC 323

502 AGACGCTTTTATAGAGGTGACATCATTAATGCGCATAGGTGTATACATCGGTATA 561  
DB 322 AGACGCTTTTATAGAGGTGACATCATTAATGCGCATAGGTGTATACATCGGTATA 263

562 CAACCTTAACCGTACACCACTTTAGCAATGAGCTGTAAATGCGCATCTCTTCGCTACCA 621  
DB 262 CAACCTTAACCGTACACCACTTTAGCAATGAGCTGTAAATGCGCATCTCTTCGCTACCA 203

622 GCACCTTTTACCAATACCTTCGCTGCTTGCAGAAACCACTGTACCAATAGCATCTACT 681  
DB 202 GCACCTTTTACCAATACCTTCGCTGCTTGCAGAAACCACTGTACCAATAGCATCTACT 143

682 GTTCT 686  
DB 142 GTTCT 138

RESULT 3  
B2314790 648 bp DNA linear GSS 06-NOV-2002  
LOCUS h22a01.b1 WGS-Zmaysef (JM107 adapted methyl filtered) Zea mays  
DEFINITION genomic clone h22a01 5', genomic survey sequence.

ACCESSION B2314790  
VERSION B2314790.1 GI:24683733  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 648)  
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,  
Katzemburger, F., King, L., Miller, B., Muller, S., Nascento, L.,  
Zutavern, T., McCombie, W.R., and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)

TITLE Lita Annenberg Hazen Genome Sequencing Center  
JOURNAL Contact: W. Richard McCombie  
COMMENT Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: h22 row: a column: 01  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 648.  
Location/Qualifiers  
1. .648  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
/clone="h22a01"

FEATURES  
source

1. .648  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
/clone="h22a01"

/lab host="JM107 or DH5a"  
/clone lib="WGS-Zmay5f (JM107 adapted methyl filtered)"  
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector  
(.x/y reads in M13mp19, .b/g reads in pUC19). The same  
ligation was transformed in either JM107 or DH5a."

## ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 648;  
Best Local Similarity 100.0%; Pred. No. 2.5e-157;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 382 ACACACTTCCACTCTGATGTTGAGTGCATCTGTTATCTCTTCCGACCATTAACAGAC 441
    |||||
DB 70 ACACACTTCCACTCTGATGTTGAGTGCATCTGTTATCTCTTCCGACCATTAACAGAC 129
    |||||

QY 442 TAGTATTTATTTGATCATGTAATGCTTATTTCTTTGAAAGCGGTTTCATTTTTTTTAC 501
    |||||
DB 130 TAGTATTTATTTGATCATGTAATGCTTATTTCTTTGAAAGCGGTTTCATTTTTTTTAC 189
    |||||

QY 502 AGACGCTTTTTTTAGAGGTGCATCATTTATGCGGCAATAGGTGTTACATGCGGTATA 561
    |||||
DB 190 AGACGCTTTTTTTAGAGGTGCATCATTTATGCGGCAATAGGTGTTACATGCGGTATA 249
    |||||

QY 562 CAACCTTAACCGTACACACATTTTGAAGATGCTGTATGCGGCAATAGGTGTTACATGCGGTATA 621
    |||||
DB 250 CAACCTTAACCGTACACACATTTTGAAGATGCTGTATGCGGCAATAGGTGTTACATGCGGTATA 309
    |||||

QY 622 GCACCTTTTACCATTAATCTTGTCTGCTGTGCAAAACCACTGTACGATAGCATCTACTGCT 681
    |||||
DB 310 GCACCTTTTACCATTAATCTTGTCTGCTGTGCAAAACCACTGTACGATAGCATCTACTGCT 369
    |||||

QY 682 GTTCT 686
    |||||
DB 370 GTTCT 374
```

RESULT 4  
B2625170/c 680 bp DNA linear GSS 16-JAN-2003  
LOCUS  
DEFINITION  
1996e05.g1 WGS-Zmay5f (DH5a methyl filtered) Zea mays genomic clone  
1996e05 5', genomic survey sequence.

ACCESSION  
B2625170  
VERSION  
B2625170.1 GI:27777030  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays

REFERENCE  
AUTHORS  
Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,  
Katzendurger, F., King, L., Miller, B., Muller, S., Nascimeto, L.,  
Zutavern, T., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)

TITLE  
JOURNAL  
COMMENT  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org

FEATURES  
SOURCE  
High quality sequence scop. 680.  
Location/Qualifiers  
1..680

/organism="Zea mays"  
/mol type="genomic DNA"  
/culivar="B73"  
/db xref="taxon:4577"  
/clone="1996e05"  
/lab host="DH5a"  
/clone lib="WGS-Zmay5f (DH5a methyl filtered)"  
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector (.x/y  
reads in M13mp19, .b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 680;  
Best Local Similarity 100.0%; Pred. No. 2.5e-157;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 382 ACACACTTCCACTCTGATGTTGAGTGCATCTGTTATCTCTTCCGACCATTAACAGAC 441
    |||||
DB 445 ACACACTTCCACTCTGATGTTGAGTGCATCTGTTATCTCTTCCGACCATTAACAGAC 386
    |||||

QY 442 TAGTATTTATTTGATCATGTAATGCTTATTTCTTTGAAAGCGGTTTCATTTTTTTTAC 501
    |||||
DB 385 TAGTATTTATTTGATCATGTAATGCTTATTTCTTTGAAAGCGGTTTCATTTTTTTTAC 326
    |||||

QY 502 AGACGCTTTTTTTAGAGGTGCATCATTTATGCGGCAATAGGTGTTACATGCGGTATA 561
    |||||
DB 325 AGACGCTTTTTTTAGAGGTGCATCATTTATGCGGCAATAGGTGTTACATGCGGTATA 266
    |||||

QY 562 CAACCTTAACCGTACACACATTTTGAAGATGCTGTATGCGGCAATAGGTGTTACATGCGGTATA 621
    |||||
DB 265 CAACCTTAACCGTACACACATTTTGAAGATGCTGTATGCGGCAATAGGTGTTACATGCGGTATA 206
    |||||

QY 622 GCACCTTTTACCATTAATCTTGTCTGCTGTGCAAAACCACTGTACGATAGCATCTACTGCT 681
    |||||
DB 205 GCACCTTTTACCATTAATCTTGTCTGCTGTGCAAAACCACTGTACGATAGCATCTACTGCT 146
    |||||

QY 682 GTTCT 686
    |||||
DB 145 GTTCT 141
```

RESULT 5  
B2320382 717 bp DNA linear GSS 06-NOV-2002  
LOCUS  
DEFINITION  
h207h12.g1 WGS-Zmay5f (JM107 adapted methyl filtered) Zea mays  
genomic clone h207h12 5', genomic survey sequence.

ACCESSION  
B2320382  
VERSION  
B2320382.1 GI:24695386  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays

REFERENCE  
AUTHORS  
Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,  
Katzendurger, F., King, L., Miller, B., Muller, S., Nascimeto, L.,  
Zutavern, T., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)

TITLE  
JOURNAL  
COMMENT  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org

FEATURES  
SOURCE  
Plate: h207 row: h column: 12



TITLE Genomic shotgun sequences from Zea mays (methyl-filtered)  
JOURNAL Unpublished (2002)  
COMMENT Contact: W. Richard McComble  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100 Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcomble@cshl.org  
Plate: h82 row: h column: 12  
Seq primer: -21M3UnivFwd  
Class: shotgun  
High quality sequence stop: 776.

FEATURES  
source  
1. .776  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="h82h12"  
/lab\_host="JM107 or DH5a"  
/note="Organ: immature ears; Site\_1: Xba I; Site\_2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector  
(x/y reads in M13mp19. b/g reads in pUC19). The same  
ligation was transformed in either JM107 or DH5a."

## ORIGIN

Query Match 25.8%; Score 305; DB 8; Length 776;  
Best Local Similarity 100.0%; Pred. No. 2.5e-157;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
382 ACACACTTCCACTGTTAGTGGATGCTCTTATCTTTTCGAAACCATTAACAGAC 441
197 ACACACTTCCACTGTTAGTGGATGCTCTTATCTTTTCGAAACCATTAACAGAC 256
442 TAGATATATTGATCATGTAATCGTTATTTCTTGAAGCGGTTTCATTTTTCAC 501
257 TAGATATATTGATCATGTAATCGTTATTTCTTGAAGCGGTTTCATTTTTCAC 316
502 AGACGCTTTTATAGAGGTGACATCCATTATGCGGCATAGGTGTATACATCCGATA 561
317 AGACGCTTTTATAGAGGTGACATCCATTATGCGGCATAGGTGTATACATCCGATA 376
562 CAACCTTAACCGTACACCATTTTATAGCAATGCGCTGTAAATGCGCATCTTCCGTAACA 621
377 CAACCTTAACCGTACACCATTTTATAGCAATGCGCTGTAAATGCGCATCTTCCGTAACA 436
622 GCACCTTTTACATTAATCTTGTCTGCTGGCAAAACCACTGTACGATAGCATTAAGCT 681
437 GCACCTTTTACATTAATCTTGTCTGCTGGCAAAACCACTGTACGATAGCATTAAGCT 496
682 GTTCT 686
497 GTTCT 501
```

RESULT 8  
CD438581/c 824 bp mRNA linear EST 03-JUN-2003  
LOCUS E010N0515A02.b Endosperm\_5 Zea mays cDNA, mRNA sequence.  
DEFINITION CD438581  
ACCESSION CD438581.1 GI:31354224  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 824)  
AUTHORS Lai J., Dey, N., Kim, C.S., Bharti, A.K., Rudd, S., Mayer, K.F.,  
Larkin, B., Beckett, P. and Messing, J.  
TITLE Characterization of the maize endosperm transcriptome and its  
JOURNAL comparison to the rice genome  
COMMENT Genome Res. 14 (10), 1932-1937 (2004)  
Contact: Lai, Jinheng  
Dr. Joachim Messing's lab  
Wakeman Institute, Rutgers University  
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
Tel: 732-445-3801  
Fax: 732-445-5735  
Email: jlai@wakeman.rutgers.edu  
Seq primer: T3

FEATURES  
source  
1. .824  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="mRNA"  
/cultivar="W22"  
/db\_xref="taxon:4577"  
/issue\_type="Endosperm of 7-23DAP"  
/clone\_lib="Endosperm\_5"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

## ORIGIN

Query Match 25.8%; Score 305; DB 6; Length 824;  
Best Local Similarity 100.0%; Pred. No. 2.5e-157;  
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
382 ACACACTTCCACTGTTAGTGGATGCTCTTATCTTTTCGAAACCATTAACAGAC 441
462 ACACACTTCCACTGTTAGTGGATGCTCTTATCTTTTCGAAACCATTAACAGAC 403
442 TAGATATATTGATCATGTAATCGTTATTTCTTGAAGCGGTTTCATTTTTCAC 501
402 TAGATATATTGATCATGTAATCGTTATTTCTTGAAGCGGTTTCATTTTTCAC 343
502 AGACGCTTTTATAGAGGTGACATCCATTATGCGGCATAGGTGTATACATCCGATA 561
342 AGACGCTTTTATAGAGGTGACATCCATTATGCGGCATAGGTGTATACATCCGATA 283
562 CAACCTTAACCGTACACCATTTTATAGCAATGCGCTGTAAATGCGCATCTTCCGTAACA 621
282 CAACCTTAACCGTACACCATTTTATAGCAATGCGCTGTAAATGCGCATCTTCCGTAACA 223
622 GCACCTTTTACATTAATCTTGTCTGCTGGCAAAACCACTGTACGATAGCATTAAGCT 681
222 GCACCTTTTACATTAATCTTGTCTGCTGGCAAAACCACTGTACGATAGCATTAAGCT 163
682 GTTCT 686
162 GTTCT 158
```

RESULT 9  
CC357906 920 bp DNA linear GSS 16-MAY-2003  
LOCUS PNHU36TD ZM 0.6.1.0 KB Zea mays genomic clone ZMMBtra454f23,  
DEFINITION genomic survey sequence.  
ACCESSION CC357906  
VERSION CC357906.1 GI:30827306  
KEYWORDS  
SOURCE  
ORGANISM  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Reinick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and  
Benutzer, J.  
TITLE Maize Genomics Consortium

## JOURNAL

Unpublished (2003)

## COMMENT

Other GSSs: FHHU36TB  
Contact: Cathy Whitelaw

## TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843Fax: 301-838-0208  
Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

## FEATURES

source

Location/Qualifiers  
1..920/organism="Zea mays"  
/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZM454P23"

/note="Vector: PCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
COT selected genomic DNA library"

## ORIGIN

Query Match

25.8%; Score 305; DB 8; Length 920;

Best Local Similarity 100.0%; Pred. No. 2.5e-157; Mismatches 0; Indels 0; Gaps 0;

Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

382 ACACACTTCCACTGAGTGTGAGTGATCGTCTGTTATCTCTCGAACCATTAACAGAC 441

303 ACACACTTCCACTGAGTGTGAGTGATCGTCTGTTATCTCTCGAACCATTAACAGAC 362

442 TAGTATTATTGATCATTGAATCGTTATTTCTCTGGAACGGTTTCATTTTTTTTAC 501

363 TAGTATTATTGATCATTGAATCGTTATTTCTCTGGAACGGTTTCATTTTTTTTAC 422

502 AGAGCTTTTGTGAGAGTGCATCATCATTAAGCGGATAGAGTGTTCATCCGCTATA 561

423 AGAGCTTTTGTGAGAGTGCATCATCATTAAGCGGATAGAGTGTTCATCCGCTATA 482

562 CAATTAACCGTACACCACTTTTGAATGAGTCTGTAATGCGGATCTCTTCGCGTACA 621

483 CAATTAACCGTACACCACTTTTGAATGAGTCTGTAATGCGGATCTCTTCGCGTACA 542

622 GCACCTTTTACATTAATCTGCTGCTGTTGCAACCCACTGTACGAATAGACTTACTCT 681

543 GCACCTTTTACATTAATCTGCTGCTGTTGCAACCCACTGTACGAATAGACTTACTCT 602

682 GTTCT 686

603 GTTCT 607

## RESULT 10

BZ375835/c

LOCUS BZ375835 600 bp DNA linear GSS 26-NOV-2002

DEFINITION 1e62D08.g1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone

ACCESSION BZ375835

VERSION BZ375835.1 GI:25463827

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 600)

AUTHORS Bhatnagar, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,

TITLE Zea mays

COMMENT Zea mays

Cold Spring Harbor Laboratory

## FEATURES

source

Location/Qualifiers  
1..600

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="1e62D08"

/lab\_host="DH5a"

/note="Organ: immature ears; Site\_1: Xba I; Site\_2: Xba I;  
The vector was digested with Xba I and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector (x/y  
reads in M13mp19). b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match

23.7%; Score 280; DB 8; Length 600;

Best Local Similarity 100.0%; Pred. No. 1.9e-143; Mismatches 0; Indels 0; Gaps 0;

Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

407 TGATCCGTTATCTCTTCGAAACCATTAACAGACTAGTATTGATCATTGAATCGT 466

600 TGATCCGTTATCTCTTCGAAACCATTAACAGACTAGTATTGATCATTGAATCGT 541

467 TTATTTCTCTGGAAGCGTTTCATTTTTTTTACAGAGCTCTTTTTTACAGAGCTCGAC 526

540 TTATTTCTCTGGAAGCGTTTCATTTTTTTTACAGAGCTCTTTTTTACAGAGCTCGAC 481

527 ATCCATTATGCGGATAGTGTATCATCATCATTAACAGACTAGTATTGATCATTGAATCGT 586

480 ATCCATTATGCGGATAGTGTATCATCATCATTAACAGACTAGTATTGATCATTGAATCGT 421

587 CAATGCGTCTGTAATGCGGATCTCTTCGCTACAGACCTTTTACCATTAATCTGCTC 646

420 CAATGCGTCTGTAATGCGGATCTCTTCGCTACAGACCTTTTACCATTAATCTGCTC 361

647 GTTGCACCAACCACTGTACGAATAGACTTACTGCTGTTCT 686

360 GTTGCACCAACCACTGTACGAATAGACTTACTGCTGTTCT 321

## RESULT 11

BZ373363

LOCUS BZ373363 600 bp DNA linear GSS 26-NOV-2002

DEFINITION 1e78F08.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone

ACCESSION BZ373363

VERSION BZ373363.1 GI:25458862

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 600)

AUTHORS Bhatnagar, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,

TITLE Zea mays

COMMENT Zea mays

Cold Spring Harbor Laboratory



Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcmombe@cehl.org  
Plate: 1678 row: F column: 08  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 622.  
Location/Qualifiers

FEATURES  
source

1. 600  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="1678E08"  
/lab\_host="DH5a"  
/clone\_11b="WGS-ZmaysF (DH5a methyl filtered)"  
/note="Organ: Immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector (x/y  
reads in M13mp19, .b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match 23.2%; Score 275; DB 8; Length 600;  
Best Local Similarity 100.0%; Pred.No.1.2e-140; Indels 0; Gaps 0;  
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 CCTGTATCTCTTCTCGAACCATTAAGACTAGTATTATTTGATCATTTGATTCGTTATT 471  
DB 1 CCTGTATCTCTTCTCGAACCATTAAGACTAGTATTATTTGATCATTTGATTCGTTATT 60  
QY 472 TCTCTTGAAGCGGTTTCATTTTCTTTTACAGAGCTCTTTTGAAGGTTCAGATCCA 531  
DB 61 TCTCTTGAAGCGGTTTCATTTTCTTTTACAGAGCTCTTTTGAAGGTTCAGATCCA 120  
QY 532 TTATGCGGCATAGGTGTTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCAATG 591  
DB 121 TTATGCGGCATAGGTGTTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCAATG 180  
QY 592 GCTGCTATAGCGGATCTCTTCCGCTACACGACCTTTTACCATTAACCTTGTCTGTGC 651  
DB 181 GCTGCTATAGCGGATCTCTTCCGCTACACGACCTTTTACCATTAACCTTGTCTGTGC 240  
QY 652 AAACCCACTGTAGCAATAGCATCTACTGCTGTCT 686  
DB 241 AAACCCACTGTAGCAATAGCATCTACTGCTGTCT 275

RESULT 12  
CC025412/c 313 bp DNA linear GSS 01-APR-2003  
LOCUS 3591.1.46.1 A05.1E1.X.1 3591 - RescueMu Grid P Zea mays genomic.  
DEFINITION genomic survey sequence.

ACCESSION CC025412.1 GI:29440269  
VERSION CC025412.1  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 313)

REFERENCE 1 Maize genomic sequences found using engineered RescueMu transposon  
AUTHORS Walbot, V.  
TITLE Unpublished (2001)  
JOURNAL Contact: Walbot V  
COMMENT Department of Biological Sciences  
Stanford University

855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 725 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Very probable ligation site of ends cut by single endonuclease.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 3591.1.46.1 row: 21  
Class: transposon-tagged.  
Location/Qualifiers

FEATURES  
source

1. 313  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73/K55"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_11b="3591 - RescueMu Grid P"  
/note="Organ: leaf; Vector: RescueMu (engineered from  
pBlueScript backbone); Site 1: BamHI; Site 2: BglII;  
RescueMu is a 4.9 kb, modified maize Mu transposon  
designed to allow plasmid rescue from total genomic DNA.  
Mu elements insert preferentially into transcription  
units. For more information on RescueMu, go to the web  
site 'www.zmdb.jastate.edu' and follow the links for  
'RescueMu.' Grid P was grown at Molokai in 2002. DNA was  
extracted from leaf strips, double digested using BamHI  
and BglII, and ligated to form circular plasmids. DH10B  
cells were transformed and then screened on LB plates with  
ampicillin."

## ORIGIN

Query Match 23.1%; Score 273; DB 8; Length 313;  
Best Local Similarity 100.0%; Pred.No.1.5e-139; Indels 0; Gaps 0;  
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 409 GATCCGTTATCTCTTCTCGAACCATTAAGACTAGTATTATTTGATCATTTGATTCGTT 468  
DB 273 GATCCGTTATCTCTTCTCGAACCATTAAGACTAGTATTATTTGATCATTTGATTCGTT 214  
QY 469 ATTTCCTTGAAGCGGTTTCATTTTCTTTTACAGAGCTCTTTTGAAGGTTCAGAT 528  
DB 213 ATTTCCTTGAAGCGGTTTCATTTTCTTTTACAGAGCTCTTTTGAAGGTTCAGAT 154  
QY 529 CCATTATGCGCATAGGTGTTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCA 588  
DB 153 CCATTATGCGCATAGGTGTTACATCGCGTATTAACCTTAACCGTACACCACTTTAGCA 94  
QY 589 ATGCGCTGTAATGCGGATCTCTTCCGCTACACGACCTTTTACCATTAACCTTGTCTGT 648  
DB 93 ATGCGCTGTAATGCGGATCTCTTCCGCTACACGACCTTTTACCATTAACCTTGTCTGT 34  
QY 649 TGCAAAACCACTGTAGCAATAGCATCTACTGCT 681  
DB 33 TGCAAAACCACTGTAGCAATAGCATCTACTGCT 1

RESULT 13  
B2359336 646 bp DNA linear GSS 18-NOV-2002  
LOCUS 1d72d02.b1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays  
DEFINITION genomic clone 1d72d02 5', genomic survey sequence.

ACCESSION B2359336.1 GI:25059030  
VERSION B2359336  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 646)

REFERENCE 1 Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,  
Katzemburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,

TITLE  
JOURNAL  
COMMENT

Zutavern, T., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: 1672 row: d column: 02  
Seq primer: -21M13univFwd  
Class: Shotgun  
High quality sequence stop: 646.  
Location/Qualifiers

FEATURES  
source

1. .646  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="1d72d02"  
/lab\_host="JM107 or DH5a"  
/note="Organ: immature ears; Site\_1: Xba I; Site\_2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector  
(.x/y reads in M13mp19, .b/g reads in pUC19). The same  
ligation was transformed in either JM107 or DH5a."

## ORIGIN

Query Match 21.6%; Score 256; DB 8; Length 646;  
Best Local Similarity 100.0%; Pred. No. 4.2e-110; Indels 0; Gaps 0;

Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ACAACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGAC 441  
DB 361 ACACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGAC 420  
QY 442 TAGATTAATTGATCATGATCGTTATTTCTCTGAAAGCGTTCAATTTTTTTTAC 501  
DB 421 TAGATTAATTGATCATGATCGTTATTTCTCTGAAAGCGTTCAATTTTTTTTAC 480  
QY 502 AGACGCTTTTGTAGAGTGCACATCATTAATGCGGATAGGTTACATCGGATTA 561  
DB 481 AGACGCTTTTGTAGAGTGCACATCATTAATGCGGATAGGTTACATCGGATTA 540  
QY 562 CAACCTTAACCGTACACCACTTTAGCAATGGCTCGTAATGCGGATCTCTTCGCTTACCA 621  
DB 541 CAACCTTAACCGTACACCACTTTAGCAATGGCTCGTAATGCGGATCTCTTCGCTTACCA 600  
QY 622 GCACCTTTTACCATTA 637  
DB 601 GCACCTTTTACCATTA 616

RESULT 14  
B2372017 650 bp DNA linear GSS 26-NOV-2002  
LOCUS B2372017  
DEFINITION 1662b08.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone  
ACCESSION B2372017  
VERSION B2372017.1 GI:25456191  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
REFERENCE 1 (bases 1 to 650)  
AUTHORS Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,

TITLE  
JOURNAL  
COMMENT

Katzenburger, F., King, L., Miller, B., Muller, S., Naecimento, L.,  
Zutavern, T., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: 1662 row: b column: 08  
Seq primer: -21M13univFwd  
Class: Shotgun  
High quality sequence stop: 651.  
Location/Qualifiers

FEATURES  
source

1. .650  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="1662b08"  
/lab\_host="DH5a"  
/note="Organ: immature ears; Site\_1: Xba I; Site\_2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector (.x/y  
reads in M13mp19, .b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match 21.4%; Score 253; DB 8; Length 650;  
Best Local Similarity 99.7%; Pred. No. 1.3e-128; Indels 1; Gaps 0;

Matches 303; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 383 CACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGACT 442  
DB 1 CACACTTCACACTAGTGTGAGTGGATCCTGTATCTCTTCGACCATTAACAGACT 60  
QY 443 AGATTAATTGATCATGATCGTTATTTCTCTGAAAGCGTTCAATTTTTTTTAC 502  
DB 61 AGATTAATTGATCATGATCGTTATTTCTCTGAAAGCGTTCAATTTTTTTTAC 480  
QY 503 GACGCTTTTGTAGAGTGCACATCATTAATGCGGATAGGTTACATCGGATTA 562  
DB 121 GACGCTTTTGTAGAGTGCACATCATTAATGCGGATAGGTTACATCGGATTA 180  
QY 563 AACTTAACCGTACACCACTTTAGCAATGGCTCGTAATGCGGATCTCTTCGCTTACCA 622  
DB 181 AACTTAACCGTACACCACTTTAGCAATGGCTCGTAATGCGGATCTCTTCGCTTACCA 240  
QY 623 CACCTTTTACCATTAATCTGCTGTTGCAAAAGCGGATGATGATATGATATGATGCTG 682  
DB 241 CACCTTTTACCATTAATCTGCTGTTGCAAAAGCGGATGATGATATGATATGATGCTG 300  
QY 683 TTCT 686  
DB 301 TTCT 304.

RESULT 15  
B2363755 428 bp DNA linear GSS 18-NOV-2002  
LOCUS B2363755  
DEFINITION 1d90d01.g1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays  
ACCESSION B2363755  
VERSION B2363755.1 GI:25068831  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
Clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 428)

TITLE  
JOURNAL  
COMMENT

Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,  
Katzemburger, F., King, L., Miller, B., Muller, S., Nascimeto, L.,  
Zutavern, T., McCombie, W.R. and Marienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished (2002)

Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: id90 row: d column: 01

Seq primer: -21M3univRev

Class: Shotgun

High quality sequence stop: 428.

FEATURES  
source

Location/Qualifiers  
1..428  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="id90d01"  
/lab\_host="JMI07 or DHSa"  
/clone\_lib="MGS-ZmaysF (JMI07 adapted methyl filtered)"  
/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 3 kb and were cloned into the vector  
(x/y reads in M13mp19, b/g reads in pUC19). The same  
ligation was transformed in either JMI07 or DHSa."

## ORIGIN

Query Match 20.4%; Score 241; DB 8; Length 428;  
Best Local Similarity 99.7%; Pred. No. 9e-122;  
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 395 CTAGTGTGAGTGATCTGTATCTTCTCGAACATAACAGACTAGTATTATTGA 454  
DB 1 CTAGTGTGAGTGATCTGTATCTTCTCGAACATAACAGACTAGTATTATTGA 60

QY 455 TCATTGAATCGTTTATTTCTCTGAAAGCGTTTCATTTTTTTTACAGAGCTTTTTT 514  
DB 61 TCATTGAATCGTTTATTTCTCTGAAAGCGTTTCATTTTTTTTACAGAGCTTTTTT 120

QY 515 TAGGAGTCGATCATTCATTATGCGGATAGGTATACATCGGATATACACTTAACCGTA 574  
DB 121 TAGGAGTCGATCATTCATTATGCGGATAGGTATACATCGGATATACACTTAACCGTA 180

QY 575 CACCACTTTTAGCAATGCTGTAAATGCGGATCTTCCGTTACAGACCTTTTACCA 634  
DB 181 CACCACTTTTAGCAATGCTGTAAATGCGGATCTTCCGTTACAGACCTTTTACCA 240

QY 635 TAACTTGTGCTGTTGCAAAACCACTGTACGAATGCAATCTACTGCTGTTCT 686  
DB 241 TAACTTGTGCTGTTGCAAAACCACTGTACGAATGCAATCTACTGCTGTTCT 292

Search completed: February 10, 2005, 17:01:49  
Job time : 8001.05 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 117.196 Seconds  
(without alignments)  
7855.666 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19  
Sequence: 1 TGTAGCGGCCACGCGTGG 19

Scoring table: OLIGO-NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description |
|------------|-------|-------------|--------|-----------|-------------|
| 1          | 19    | 100.0       | 19     | AX342370  | Sequence    |
| 2          | 19    | 100.0       | 498    | AX342368  | Sequence    |
| 3          | 18    | 94.7        | 3178   | AX492266  | Corynebact  |
| 4          | 16    | 84.2        | 189    | AY271325  | Mycobacte   |
| 5          | 16    | 84.2        | 189    | AY271340  | Mycobacte   |
| 6          | 16    | 84.2        | 204    | AY271897  | Mycobacte   |
| 7          | 16    | 84.2        | 3180   | AY492278  | Corynebact  |
| 8          | 16    | 84.2        | 3330   | AY492261  | Corynebact  |
| 9          | 16    | 84.2        | 3333   | AY492259  | Corynebact  |
| 10         | 16    | 84.2        | 3357   | AY492234  | Corynebact  |
| 11         | 16    | 84.2        | 3357   | AY492285  | Rhodococc   |
| 12         | 16    | 84.2        | 3702   | AY1262743 | Mycobacte   |
| 13         | 16    | 84.2        | 3868   | AY147167  | Mycobacte   |
| 14         | 16    | 84.2        | 3906   | AY147173  | Mycobacte   |
| 15         | 16    | 84.2        | 3929   | AY262736  | Mycobacte   |
| 16         | 16    | 84.2        | 4405   | AY147165  | Mycobacte   |
| 17         | 16    | 84.2        | 299050 | SC0939104 | Streptomy   |
| 18         | 16    | 84.2        | 301675 | AP005027  | Streptomy   |
| 19         | 16    | 84.2        | 316296 | AC092285  | Homo sapi   |

|      |    |      |       |    |            |                     |
|------|----|------|-------|----|------------|---------------------|
| C 20 | 15 | 78.9 | 1183  | 10 | RATREP3B   | D16443 Rattus norv  |
| C 21 | 15 | 78.9 | 1211  | 10 | RNEP3B     | X80133 R.norvegicu  |
| C 22 | 15 | 78.9 | 1253  | 10 | RATREP3R   | D29969 Rat mRNA fo  |
| C 23 | 15 | 78.9 | 1275  | 10 | AF302686   | AF302686 Rattus no  |
| C 24 | 15 | 78.9 | 1310  | 10 | RATREP3    | D14869 Rattus norv  |
| C 25 | 15 | 78.9 | 1345  | 10 | RNEP3ALPH  | X83855 R.norvegicu  |
| C 26 | 15 | 78.9 | 1405  | 6  | AR038125   | AR038125 Sequence   |
| C 27 | 15 | 78.9 | 1405  | 6  | I28939     | I28939 Sequence 3   |
| C 28 | 15 | 78.9 | 1616  | 8  | AK111750   | AK111750 Oryza sat  |
| C 29 | 15 | 78.9 | 1974  | 4  | AF266477   | AF266477 Canis fam  |
| C 30 | 15 | 78.9 | 2002  | 10 | MUSPGERB   | D13321 Mus musculu  |
| C 31 | 15 | 78.9 | 2006  | 6  | E06015     | E06015 DNA encodin  |
| C 32 | 15 | 78.9 | 2090  | 10 | BC058742   | BC058742 Mus muscu  |
| C 33 | 15 | 78.9 | 2091  | 10 | MUSPGER    | D10204 Mus musculu  |
| C 34 | 15 | 78.9 | 2107  | 6  | AR038124   | AR038124 Sequence   |
| C 35 | 15 | 78.9 | 2107  | 6  | E06014     | E06014 DNA encodin  |
| C 36 | 15 | 78.9 | 2107  | 6  | I28938     | I28938 Sequence 1   |
| C 37 | 15 | 78.9 | 2556  | 14 | HSVTXY     | K02122 Matmoseet he |
| C 38 | 15 | 78.9 | 2811  | 5  | GGU47273   | U47273 Gallus gall  |
| C 39 | 15 | 78.9 | 3245  | 5  | CHKTRANGLU | L02270 Chicken tra  |
| C 40 | 15 | 78.9 | 3570  | 10 | MUSPGERB   | D17406 Mus musculu  |
| C 41 | 15 | 78.9 | 6219  | 1  | AF340167   | AF340167 Streptomy  |
| C 42 | 15 | 78.9 | 7000  | 10 | AF458960   | AF458960 Mus muscu  |
| C 43 | 15 | 78.9 | 15674 | 9  | AC118277   | AC118277 Homo sapi  |
| C 44 | 15 | 78.9 | 51855 | 1  | ML1557546  | AJ557546 Melittang  |
| C 45 | 15 | 78.9 | 52684 | 8  | AF111709   | AF111709 Oryza sat  |

#### ALIGNMENTS

|                       |            |  |               |            |         |                 |
|-----------------------|------------|--|---------------|------------|---------|-----------------|
| RESULT 1              | AX342370   | Sequence 9 from Patent EP1167531.                                | 19 bp         | DNA        | linear  | PAT 12-JAN-2002 |
| LOCUS                 | AX342370   |  |               |            |         |                 |
| DEFINITION            | AX342370   |  |               |            |         |                 |
| ACCESSION             | AX342370.1 | GI:18151813  |               |            |         |                 |
| VERSION               |            |  |               |            |         |                 |
| KEYWORDS              |            |  |               |            |         |                 |
| SOURCE                |            |  |               |            |         |                 |
| ORGANISM              |            |  |               |            |         |                 |
| REFERENCE             | 1          | Behr, C.F., Hironaka, C., Heck, G.R. and You, J.                 |               |            |         |                 |
| AUTHORS               |            | Corn transformant pv-zmgf32 (nke03) and compositions and methods |               |            |         |                 |
| TITLE                 |            | for detection thereof  |               |            |         |                 |
| JOURNAL               |            | Patent: EP 1167531-A 9 02-JAN-2002;                              |               |            |         |                 |
| FEATURES              |            | Monsanto Technology LLC (US)                                     |               |            |         |                 |
| SOURCE                |            | Location/Qualifiers  |               |            |         |                 |
|                       |            | 1..19  |               |            |         |                 |
|                       |            | /organism="synthetic construct"                                  |               |            |         |                 |
|                       |            | /mol_type="unassigned DNA"                                       |               |            |         |                 |
|                       |            | /db_xref="taxon:32630"   |               |            |         |                 |
|                       |            | /note="Zea maize genomic and vector DNA"                         |               |            |         |                 |
| ORIGIN                |            |  |               |            |         |                 |
| Query Match           | 100.0%     | Score 19;  | DB 6;         | Length 19; |         |                 |
| Best Local Similarity | 100.0%     | Pred. No. 24;  |               |            |         |                 |
| Matches               | 19;        | Conservative 0;  | Mismatches 0; | Indels 0;  | Gaps 0; |                 |
| Cy                    | 1          | TGTAGCGGCCACGCGTGG   | 19            |            |         |                 |
| Db                    | 1          | TGTAGCGGCCACGCGTGG   | 19            |            |         |                 |
| RESULT 2              | AX342368   | Sequence 7 from Patent EP1167531.                                | 498 bp        | DNA        | linear  | PAT 12-JAN-2002 |
| LOCUS                 | AX342368   |  |               |            |         |                 |
| DEFINITION            | AX342368   |  |               |            |         |                 |
| ACCESSION             | AX342368.1 | GI:18151811  |               |            |         |                 |
| VERSION               |            |  |               |            |         |                 |
| KEYWORDS              |            |  |               |            |         |                 |
| SOURCE                |            |  |               |            |         |                 |
| ORGANISM              |            |  |               |            |         |                 |
|                       |            | synthetic construct  |               |            |         |                 |
|                       |            | synthetic construct  |               |            |         |                 |

other sequences; artificial sequences.

REFERENCE 1  
Behr, C.F., Hironaka, C., Heck, G.R. and You, J.  
Corn transformant PV-zmgf32 (nk603) and compositions and methods  
for detection thereof  
Patent: EP 1167531-A 7 02-JAN-2002;  
Monsanto Technology LLC (US)  
JOURNAL location/Qualifiers  
source 1. 498  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="1-304 Zea maize genomic DNA 305-349 construct  
vector DNA 350-498 rice actin 1 promoter DNA"

## ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 498;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGCGCCACGCGTGG 19  
|||||  
Db 296 TGTAGCGCGCCACGCGTGG 314

RESULT 3  
AY492266/c 3178 bp DNA linear BCT 14-SEP-2004  
LOCUS Corynebacterium afermentans subsp. lipophilum strain CIP 103500  
DEFINITION RPOB (rpoB) gene, partial cds..  
ACCESSION AY492266  
VERSION AY492266.1 GI:46360989  
KEYWORDS  
SOURCE Corynebacterium afermentans subsp. lipophilum  
ORGANISM Corynebacterium afermentans subsp. lipophilum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1 (bases 1 to 3178)  
AUTHORS Khamis, A., Raoult, D. and La Scola, B.  
TITLE rpoB Gene Sequencing for Identification of Corynebacterium Species  
JOURNAL J. Clin. Microbiol. 42 (9), 3925-3931 (2004)  
PUBMED 15364970  
REFERENCE 2 (bases 1 to 3178)  
AUTHORS La Scola, B., Khamis, A. and Raoult, D.  
TITLE Direct Submission  
JOURNAL Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,  
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille  
13385, France

## FEATURES

source

1. 3178  
/organism="Corynebacterium afermentans subsp. lipophilum"  
/mol\_type="genomic DNA"  
/strain="CIP 103500"  
/sub\_species="lipophilum"  
/db\_xref="taxon:144184"  
/clone="carpob37"  
/note="type strain of Corynebacterium afermentans subsp.  
lipophilum"  
1. .>3178  
/gene="rpoB"  
<1. .>3178  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=11  
/product="RpoB"  
/protein\_id="AAS89228.1"  
/db\_xref="GI:46360990"  
/translation="TAAVSDOTMNAADPGAPERYSPAKINEPITVPGLLDVOLSPFAW  
LVGQEMREERBOANRGDARTSGLEDILIEISIDYSGMSLTSPREEDVKYTI  
DECKDKINYSAPLYVTAFINDDTQIKSTVITGPFPLMTDKGTFTVNSTEKVYVS  
QIVRSFGVYDEITDKSTERPLSHVKVIPSAGAWLEFVDVDRDTVGVAIDKKRQPVAT  
VLKALGWTTEQITERGFSFISIMSTLENDVSNIDEALLEIYKQRFGEQPTDLAQ  
SLIENSFFKAKRYDLARVGRYKTKIGLGDHDLMTLTEDTATLTLEYLVRLHAG

gene  
CDS

## ORIGIN

Query Match 94.7%; Score 18; DB 1; Length 3178;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGCGCCACGCGTGG 18  
|||||  
Db 892 TGTAGCGCGCCACGCGTGG 875

RESULT 4  
AY271325/c 189 bp DNA linear BCT 07-MAY-2003  
LOCUS Mycobacterium fortuitum type I RNA polymerase beta subunit (rpoB)  
DEFINITION gene, partial cds..  
ACCESSION AY271325  
VERSION AY271325.1 GI:30145567  
KEYWORDS  
SOURCE Mycobacterium fortuitum  
ORGANISM Mycobacterium fortuitum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 189)  
AUTHORS Lee, H., Bang, H. E., Bai, G. H. and Cho, S. N.  
TITLE Novel Polymorphic Region of the rpoB Gene Containing Mycobacterium  
Species-Specific Sequences and Its Use in Identification of  
Mycobacteria  
JOURNAL J. Clin. Microbiol. 41 (5), 2213-2218 (2003)  
PUBMED 12734283  
REFERENCE 2 (bases 1 to 189)  
AUTHORS Bang, H.-E., Park, H.-J., Bai, G.-H., Cho, S.-N. and Lee, H.  
TITLE Direct Submission  
JOURNAL Submitted (03-APR-2003) Biomedical Laboratory Science, College of  
Health Sciences, Maeji-Ri, Heungup-Myun, Wonju, Kangwon-do 220-710,  
Korea

## FEATURES

source

1. 189  
/organism="Mycobacterium fortuitum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1766"  
/note="type: I"  
<1. .>189  
/gene="rpoB"  
<1. .>189  
/gene="rpoB"  
/codon\_start=2  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="AAP15214.1"  
/db\_xref="GI:30145568"  
/translation="LARVGRYKVNKKGLGNAGOPTTSITLTEDVATTEYLVRLHAG  
QTWTVGVGEVEVD"

gene  
CDS

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 189;  
Best Local Similarity 100.0%; Pred. No. 4; 4e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGCGCCACGCG 16

Db 23 TGTAGCGGCCACGCG 8

RESULT 5  
LOCUS AY271340/c 189 bp DNA linear BCT 07-MAY-2003  
DEFINITION Mycobacterium fortuitum type II RNA polymerase beta subunit (rpoB)  
ACCESSION AY271340  
VERSION AY271340  
KEYWORDS gene, partial cds.  
SOURCE AY271340.1 GI:30145597  
ORGANISM Mycobacterium fortuitum  
Mycobacterium fortuitum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 189)  
AUTHORS Lee, H., Bang, H. E., Bai, G. H. and Cho, S. N.  
TITLE Novel Polymorphic Region of the rpoB Gene Containing Mycobacterium  
Species-Specific Sequences and Its Use in Identification of  
Mycobacteria  
J. Clin. Microbiol. 41 (5), 2213-2218 (2003)  
12734283  
PUBMED 12734283  
REFERENCE 2 (bases 1 to 189)  
AUTHORS Bang, H. E., Park, H. J., Bai, G. H., Cho, S. N. and Lee, H.  
TITLE Direct Submission  
Submitted (03-APR-2003) Biomedical Laboratory Science, College of  
Health Sciences, Maeji-Ri, Heungup-Myun, Wonju, Kangwon-Do 220-710,  
Korea

FEATURES  
source location/Qualifiers  
1..189  
/organism="Mycobacterium fortuitum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1766"  
/note="type: II"  
<1..>189  
/gene="rpoB"  
<1..>189  
/gene="rpoB"  
/codon\_start=2  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="AAPI5229.1"  
/db\_xref="GI:30145598"  
/translation="LAVRGYKVKKLGILNAGDPITSTLTEDVATIEYLRLHMG  
QTMVTPGVGPVEVVD"

ORIGIN  
Query Match 84.2%; Score 16; DB 1; Length 189;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCG 16  
|||||  
23 TGTAGCGGCCACGCG 8

Db 23 TGTAGCGGCCACGCG 8

RESULT 6  
LOCUS AY271897/c 204 bp DNA linear BCT 07-MAY-2003  
DEFINITION Mycobacterium kansasii type III RNA polymerase beta subunit (rpoB)  
ACCESSION AY271897  
VERSION AY271897  
KEYWORDS gene, partial cds.  
SOURCE AY271897.1 GI:30145609  
ORGANISM Mycobacterium kansasii  
Mycobacterium kansasii  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 204)  
AUTHORS Lee, H., Bang, H. E., Bai, G. H. and Cho, S. N.  
TITLE Novel Polymorphic Region of the rpoB Gene Containing Mycobacterium  
Species-Specific Sequences and Its Use in Identification of

Myobacteria  
J. Clin. Microbiol. 41 (5), 2213-2218 (2003)  
12734283  
PUBMED 12734283  
REFERENCE 2 (bases 1 to 204)  
AUTHORS Bang, H. E., Park, H. J., Bai, G. H., Cho, S. N. and Lee, H.  
TITLE Direct Submission  
Submitted (10-APR-2003) Biomedical Laboratory Science, College of  
Health Sciences, Maeji-Ri, Heungup-Myun, Wonju, Kangwon-Do 220-710,  
Korea

FEATURES  
source location/Qualifiers  
1..204  
/organism="Mycobacterium kansasii"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1768"  
/note="type: III"  
<1..>204  
/gene="rpoB"  
<1..>204  
/gene="rpoB"  
/codon\_start=2  
/transl\_table=11  
/product="RNA polymerase beta subunit"  
/protein\_id="AAPI5235.1"  
/db\_xref="GI:30145610"  
/translation="LAVRGYKVKKLGILNAGDPITSTLTEDVATIEYLRLHHA  
SDQGPANWTPGVGPVEVETDD"

ORIGIN  
Query Match 84.2%; Score 16; DB 1; Length 204;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCG 16  
|||||  
23 TGTAGCGGCCACGCG 8

Db 23 TGTAGCGGCCACGCG 8

RESULT 7  
LOCUS AY492278/c 3180 bp DNA linear BCT 14-SEP-2004  
DEFINITION Corynebacterium riegellii strain CIP 105310 rpoB (rpoB) gene,  
partial cds.  
ACCESSION AY492278  
VERSION AY492278  
KEYWORDS  
SOURCE AY492278.1 GI:46361013  
ORGANISM Corynebacterium riegellii  
Corynebacterium riegellii  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE 1 (bases 1 to 3180)  
AUTHORS Khamis, A., Raoult, D. and La Scola, B.  
TITLE rpoB Gene Sequencing for Identification of Corynebacterium Species  
J. Clin. Microbiol. 42 (9), 3925-3931 (2004)  
15364970  
PUBMED 15364970  
REFERENCE 2 (bases 1 to 3180)  
AUTHORS La Scola, B., Khamis, A. and Raoult, D.  
TITLE Direct Submission  
Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,  
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille  
13385, France

FEATURES  
source location/Qualifiers  
1..3180  
/organism="Corynebacterium riegellii"  
/mol\_type="genomic DNA"  
/strain="CIP 105310"  
/db\_xref="taxon:156976"  
/clone="scrrpob49"  
/note="type strain of Corynebacterium riegellii"  
<1..>3180  
/gene="rpoB"  
<1..>3180  
/gene="rpoB"  
/codon\_start=1

/transl\_table=11  
/product="RpoB"  
/protein\_id="AA892240.1"  
/db\_xref="GI:46361014"  
/translation="LAVSDQTMMAEIPGAPERYSPAKINEPIVPGLLDVOLVESFAM  
LVGTSEEREQOQRGFTARVTSGLIEDILEISPIQDYSGNMSTLSSEPRDEVKTYI  
EAKDKDINYSAPLYTAERINNDTOEIKSQTVIIGDPLMTEKGTPIVNGTERVVS  
QVRSFGVDEFTIDKSTERPLHSVKYIPSKGAMLEFDVDRDVGVRIDKRRQPT  
VLKALGWTQEITRETFSEIMMAITLESQVANTDALLEIYKQKQGEPTDLAQ  
SLENSFFKAKRQDLARVGRYKVRKLGSDHGLMTLEEDIAATLEYVRLHADE  
SEMTSPSGEIIPISTDDIDHGRRLRPTNGELLONQVGLSMERYVREMTTODE  
SITPISLINRVSAAITREFGTSLSQFMDONNSLSGLTHKRRLSLQKNDKLTLD  
GIEVDVPSHYGMCPIETPEGNIGLIGALASVAVNPPGFIETPYQKNDKLTLD  
QVYLTADBEDRYAIAQAATPMDKGNLTGERIEVRLKDGDIYVVGPKGYDIDSPR  
QMSVATAMIPLEHDDANRALMGANMOQAAPLRAAAYATGMEORAAVYDADTV  
ISPKAGVETVGTGIVTMDDEGRDTYMLTERTNOGTCYNOTPIVSOGDRVEAGQ  
VIADGCTKGEVMAKRNLLVAPNPMGNYEDAIINQRYVEDILTSHYIEHEED  
ARDTKGAEITREIPNVSEVLDKDLBERGIRIGADVRDSDIIVGKTYRGETELP  
EERLLRAIFEEKAREVDTSLKVPBGQKVIARREARDEDDDLARGVEMIRVYA  
QKRIKIDGDMAGHGNKGVKILPOEDMFMADGTPVDIILNTHGPRMNIQVYA  
EVLHGLAKAGMTVNPDPKNAKLETLPEHLVYDPAVNSLTATPFDGATNDEIAGLL  
ANSKPNRSDGVMDENGKTMLEPDRSGSEPYKPIISVGYMTMLKHLHLDKRIHARSIG  
PYSMITQOPLGKRAQFGGQRFGEHEWMAQ"

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3180;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16  
|||||  
Db 892 TGTAGCGGCCCGCG 877

RESULT 8 3330 bp DNA linear BCT 14-SEP-2004  
AY492261/c  
LOCUS  
DEFINITION Corynebacterium mucifaciens strain CIP 105129 RpoB (rpoB) gene, partial cds.  
ACCESSION AY492261  
VERSION AY492261.1 GI:46360979  
KEYWORDS  
SOURCE  
ORGANISM Corynebacterium mucifaciens  
Corynebacterium mucifaciens  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE  
AUTHORS 1 (bases 1 to 3330)  
TITLE Khamis, A., Raoult, D. and La Scola, B.  
JOURNAL rpoB Gene Sequencing for Identification of Corynebacterium Species  
PUBMED 15364970  
AUTHORS 2 (bases 1 to 3330)  
TITLE La Scola, B., Khamis, A. and Raoult, D.  
JOURNAL Direct Submission  
PUBMED  
TITLE Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,  
JOURNAL Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille  
13385, France  
FEATURES  
source  
Location/Qualifiers  
1..3330  
/organism="Corynebacterium mucifaciens"  
/mol\_type="genomic DNA"  
/strain="CIP 105129"  
/db\_xref="taxon:57171"  
/clone="cmrpb32"  
/note="type strain of Corynebacterium mucifaciens"  
1..3330  
/gene="rpoB"  
<1..3330  
/codon\_start=3  
/transl\_table=11  
/product="RpoB"  
/protein\_id="AA892240.1"

/db\_xref="GI:46360980"  
/translation="SRQTMMAEIPGAPERYSPAKINEPIVPGLLDVOLVESFAMLVG  
TPMERREQAMSGDARITSGLEDIDEISPIEDYSGNMSLTLSSEPRDEVKTYIDEC  
KDKDINYSAPLYTAERINNDTOEIKSQTVIIGDPLMTEKGTPIVNGTERVVSQV  
RSFGVDEFTIDKSTERPLHSVKYIPSKGAMLEFDVDRDVGVRIDKRRQPTVL  
KALGWTQEITRETFSEIMMAITLESQVANTDALLEIYKQKQGEPTDLAQSL  
ENSFFAKRDILARVGRYKVRKLGSDHGLMTLEEDIAATLEYVRLHADESEM  
TSPEGELIPIINTDDIDHGRRLRPTNGELLONQVGLSMERYVREMTTODESIT  
PTILINRVSAAITREFGTSLSQFMDONNSLSGLTHKRRLSLQKNDKLTLD  
VRSVSHYGRMCPIETPEGNIGLIGALASVAVNPPGFIETPYQKNDKLTLD  
YLTADBEDRYAIAQAATPMDKGNLTGERIEVRLKDGDIYVVGPKGYDIDSPR  
QMSVATAMIPLEHDDANRALMGANMOQAAPLRAAAYATGMEORAAVYDADTV  
ISPKAGVETVGTGIVTMDDEGRDTYMLTERTNOGTCYNOTPIVSOGDRVEAGQ  
VIADGCTKGEVMAKRNLLVAPNPMGNYEDAIINQRYVEDILTSHYIEHEED  
ARDTKGAEITREIPNVSEVLDKDLBERGIRIGADVRDSDIIVGKTYRGETELP  
EERLLRAIFEEKAREVDTSLKVPBGQKVIARREARDEDDDLARGVEMIRVYA  
QKRIKIDGDMAGHGNKGVKILPOEDMFMADGTPVDIILNTHGPRMNIQVYA  
EVLHGLAKAGMTVNPDPKNAKLETLPEHLVYDPAVNSLTATPFDGATNDEIAGLL  
ANSKPNRSDGVMDENGKTMLEPDRSGSEPYKPIISVGYMTMLKHLHLDKRIHARSIG  
PYSMITQOPLGKRAQFGGQRFGEHEWMAQ

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3330;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16  
|||||  
Db 885 TGTAGCGGCCCGCG 870

RESULT 9 3333 bp DNA linear BCT 14-SEP-2004  
AY492259/c  
LOCUS  
DEFINITION Corynebacterium imitans strain CIP 105130 RpoB (rpoB) gene, partial cds.  
ACCESSION AY492259  
VERSION AY492259.1 GI:46360975  
KEYWORDS  
SOURCE  
ORGANISM Corynebacterium imitans  
Corynebacterium imitans  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
REFERENCE  
AUTHORS 1 (bases 1 to 3333)  
TITLE Khamis, A., Raoult, D. and La Scola, B.  
JOURNAL rpoB Gene Sequencing for Identification of Corynebacterium Species  
PUBMED 15364970  
AUTHORS 2 (bases 1 to 3333)  
TITLE La Scola, B., Khamis, A. and Raoult, D.  
JOURNAL Direct Submission  
PUBMED  
TITLE Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR6020 IFR48,  
JOURNAL Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille  
13385, France  
FEATURES  
source  
Location/Qualifiers  
1..3333  
/organism="Corynebacterium imitans"  
/mol\_type="genomic DNA"  
/strain="CIP 105130"  
/db\_xref="taxon:156978"  
/clone="cmrpb30"  
/note="type strain of Corynebacterium imitans"  
1..3333  
/gene="rpoB"  
<1..3333  
/codon\_start=3  
/transl\_table=11  
/product="RpoB"  
/protein\_id="AA892241.1"  
/db\_xref="GI:46360976"  
/translation="SRQTMMAEIPGAPERYSPAKINEPIVPGLLDVOLVESFAMLVG



TEWERBOELRGDEARVKSGLIEDIDEISPIODYSGNMSLTISEPREFDAKTYTEEA  
KODINVSAPLYVTAEPFINNDQIEKQTVF1GDFPLMTDKGTFVINGTERVVSQV  
RSGVYFDEITDKSTERPLHAYVIPSRGAMLEFDVKRDVGRV1DRRRORPVTYLL  
KALGWTBOITERFGFSELMSTLENDGVAMTEDEALLEIYRKORPEGEPTDIAOSIL  
ENSPFAKRYDVARVGRVKNRKLGLGSDHEDLTTEED1ATTELYLVRHAGETEM  
TSPSGGIIITINDDD1DHFGNRLRTVGEILQONOVVGLSRMEVNEREMTDOASIT  
PSTLINVRPVSAAIRREFGTSQSDPMDDNNSLSGTHRRISALCPGLSRRAGIE  
VADVHPSHGRMCPIETEPGPNIGLIGALSAYARVNPFG1ETPYKVDGRLTDQID  
YLTADEEDRYAIAQAATPMTKOGELTGER1EVRKXGDIGVGPQGVADYDLSRQW  
SVATAMIPILEHDDNARALMAGNMKQAVLEAARVATGMEQRAAYDAGDTVTR  
LGVATVTVGDTYTVAGDGTDRPTQOLRTPHRTNOGTCYNOVPIVSAQORVEAGVIA  
DEPGTNGMSLGRN1LVAFMPEBGNRYDA11INORVYBED1LTSVH1EEHEDARD  
TKLGAELIRE1PNVSEDLKOLDEKGI1RIGADVDDG1LVKVTPKGETELTPPER  
L1RA1FGEKAREVRDTS1KVPHEGTSKV1AVERFSREDD1LPGVNM1R1VVAQK  
KIODGKMGARHNGKGVGK1LPEDMPFMEDEGT1PVDI1LNTHGVPRRNM1GQVLEH  
LGM1AKAGMTVNDPDPANAME1LTPLEKLYDVPES1LATVPFGASNE1AG1LLANS  
KPRDGDVWVNGDGKAR1LDGSGSPKPYK1SVGYMYMLKH1L1VDEKIHARS1GTPS  
MITQPLGKAQGGORFGEMERMAQAYGAAT1TLOEL1T1KSDDVVGRVYEA1VK  
GDN1PDGP1PESFKVLKELQ1S

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 333;

Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCCCACGCG 16  
|||||  
Db 885 TGTAGCGGCCCCACGCG 870

RESULT 10  
AY492234/c 3357 bp DNA linear BCT 14-SEP-2004  
LOCUS Corynebacterium auris strain CIP 104632 RpoB (rpoB) gene, partial  
DEFINITION

ACCESSION  
AY492234  
VERSION  
AY492234.1 GI:46360925

SOURCE  
ORGANISM  
Corynebacterium auris  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE  
AUTHORS  
Khamis, A., Raoult, D. and La Scola, B.  
TITLE  
rpoB Gene Sequencing for Identification of Corynebacterium Species

JOURNAL  
J Clin. Microbiol. 42 (9), 3925-3931 (2004)

PUBMED  
15364970  
2 (bases 1 to 3357)

REFERENCE  
AUTHORS  
La Scola, B., Khamis, A. and Raoult, D.  
TITLE  
Direct Submission  
Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR620 IFR48,  
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille  
13385, France

FEATURES  
source  
Location/Qualifiers

1..3357  
/organism="Corynebacterium auris"  
/mol\_type="genomic DNA"  
/strain="CIP 104632"  
/db\_xref="taxon:44750"  
/clone="carp05"  
/note="type strain of Corynebacterium auris"  
<1..>3357  
/gene="rpoB"  
<1..>3357  
/gene="rpoB"  
/codon\_start=3  
/transl\_table=11  
/product="rpoB"  
/protein\_id="AAS89196.1"  
/db\_xref="GI:46360926"  
/translation="SRQMSMATTPGAPORVSPAKISRPITVPGLLDVORSPFVAVG  
TPEWERQARGBGVRTSGIEDILELSP1ODYSGNMSLS1SPREDYKTYTIDEC  
KODINVSAPLYVTAEPFINNDQIEKQTVF1GDFPLMTDKGTFVINGTERVVSQV

RSGVYFDEITDKSTERPLHAYVIPSRGAMLEFDVKRDVGRV1DRRRORPVTYLL  
KALGWTBOITERFGFSELMSTLESDGVAMTEDEALLEIYRKORPEGEPTDIAOSIL  
ENAPFRARV1DARVGRVKNRKLGLGSDHEDLTTEED1ATTELYLVRHAGETEM  
TSPNGAVATINDDD1DHFGNRLRTVGEILQONOVVGLSRMEVNEREMTDOASIT  
PSTLINVRPVSAAIRREFGTSQSDPMDDNNSLSGTHRRISALCPGLSRRAGIE  
VADVHPSHGRMCPIETEPGPNIGLIGALSAYARVNPFG1ETPYKVDGRLTDQID  
YLTADEEDRYAIAQAATPMTKOGELTGER1EVRKXGDIGVGPQGVADYDLSRQW  
SVATAMIPILEHDDNARALMAGNMKQAVLEAARVATGMEQRAAYDAGDTVTR  
LGVATVTVGDTYTVAGDGTDRPTQOLRTPHRTNOGTCYNOVPIVSAQORVEAGVIA  
DEPGTNGMSLGRN1LVAFMPEBGNRYDA11INORVYBED1LTSVH1EEHEDARD  
TKLGAELIRE1PNVSEDLKOLDEKGI1RIGADVDDG1LVKVTPKGETELTPPER  
L1RA1FGEKAREVRDTS1KVPHEGTSKV1AVERFSREDD1LPGVNM1R1VVAQK  
KIODGKMGARHNGKGVGK1LPEDMPFMEDEGT1PVDI1LNTHGVPRRNM1GQVLEH  
LGM1AKAGMTVNDPDPANAME1LTPLEKLYDVPES1LATVPFGASNE1AG1LLANS  
KPRDGDVWVNGDGKAR1LDGSGSPKPYK1SVGYMYMLKH1L1VDEKIHARS1GTPS  
MITQPLGKAQGGORFGEMERMAQAYGAAT1TLOEL1T1KSDDVVGRVYEA1VK  
GDN1PDGP1PESFKVLKELQ1SCLNVEVL"

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3357;

Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCCCACGCG 16  
|||||  
Db 885 TGTAGCGGCCCCACGCG 870

RESULT 11  
AY492285 3357 bp DNA linear BCT 14-SEP-2004  
LOCUS Rhodococcus equi strain CIP 81.17 RpoB (rpoB) gene, partial cds.  
DEFINITION  
ACCSSION  
AY492285  
VERSION  
AY492285.1 GI:46361027

SOURCE  
ORGANISM  
Rhodococcus equi  
Rhodococcus equi

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Nocardiaceae; Rhodococcus.

REFERENCE  
AUTHORS  
Khamis, A., Raoult, D. and La Scola, B.  
TITLE  
rpoB Gene Sequencing for Identification of Corynebacterium Species

JOURNAL  
J Clin. Microbiol. 42 (9), 3925-3931 (2004)

PUBMED  
15364970  
2 (bases 1 to 3357)

REFERENCE  
AUTHORS  
La Scola, B., Khamis, A. and Raoult, D.  
TITLE  
Direct Submission  
Submitted (01-DEC-2003) Unite des Rickettsies CNRS UMR620 IFR48,  
Faculte de Medecine de Marseille, 27 Bd. Jean Moulin, Marseille  
13385, France

FEATURES  
source  
Location/Qualifiers

1..3357  
/organism="Rhodococcus equi"  
/mol\_type="genomic DNA"  
/strain="CIP 81.17"  
/db\_xref="taxon:43767"  
/clone="rep0056"  
/note="conspecificity: Corynebacterium equi (type strain  
CIP 54.72) = Corynebacterium hoagii (type strain CIP  
81.17); type strain of Corynebacterium hoagii"  
<1..>3357  
/gene="rpoB"  
<1..>3357  
/gene="rpoB"  
/codon\_start=3  
/transl\_table=11  
/product="rpoB"  
/protein\_id="AAS89247.1"  
/db\_xref="GI:46361028"  
/translation="SRQTKAVAGIIPGAPRRVSPAKIRPLEVPGLLDQTSFENVG  
SPSWREBAARAGVAGVGLIEILAEISPIEDFSGMSLSFSDFBEVKSVECKD  
KDMTYAPLFTVTAEPFINNDQIEKQTVF1GDFPLMTDKGTFVINGTERVVSQVRS  
PGVYFDSIDSTKTHLSAVV1IGRGAMLEFDVKDVTGVR1DRRRORPVTYLLKA

LGMTDEIRERFGFSEIMWATLEKDPKANTDEALDIYRKLRGPSEPPTKSAQTLLN  
LPEKDKRYDLARVGRYKINKKLGIATGLPIEASTLTEDDITVITTEYLRVLRAGTMMT  
APGVEPVVVDIDHFGNRRRLRTVGLIQLVIRVGLSMEVREEMTDDVATIP  
OTLINRPPVAAIKIEFPGTSQLSQFOMDONNLSGLTHRRRLSALGPGSLSERGLV  
RDVPSHYGMCPIETEPGPNIGIGSLSVARAPPGCTIETTPRYKINQOLTPONVY  
LTDEEDRRHVAAQANSPYDANGRTERTERVLRKRGSEVERFSSDVMYDSPPQMS  
VAAVAMIPELHDDANRALMGANNMORQAVPLVRSEAPLVGTMEILAAVADAVLTER  
TGVAEVSDDYVYVMADGSRRTYRLKRFANNOGTANORPYDEGORYAGOVLD  
GPTCEENEMALGKULVAIMPBEHNYEDAIILSQVVEEDVLSIHIEHEIDARDT  
KLGAETITRDIPIVNSDEVADLDLBERGIVRIGAEVARDGILVAGKTPKGETLTPBERL  
LRAIFGKAREBDTSLKVPBGSGKIVIGRVSFREDDDLPQGVNELVRYVQOKK  
IODGDKLAGHNGKGVTKILPQDMRPLPDGTVDIILNTHGVPRRMNIGQVTEHL  
GMIGKTMVQIAGDSRPFDAATLPEEMLSAPASNSINATVFDGAKEDLTGLGST  
LPRNDEGRWVPPDGKATLFDGSGEPPEPVSVGMVYIKLHLVDKIHARSSTGPS  
MITOOLPGKAGOFGGORFGMECWMQAYGAAYTLOELITIKSDIVVGRVAYEAIYK  
GENIPERGISFVILKELQSLCLNVEVL"

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3357;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCG 16

Db 879 TGTAGCGGCCACGCG 864

## RESULT 12

AY262743/c

LOCUS 3702 bp DNA linear BCT 13-JAN-2004  
DEFINITION Mycobacterium wolinskyi strain ATCC 700010 RpoB gene, complete cds.

AY262743

VERSION

AY262743.1 GI:34595758

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

14662964

2 (bases 1 to 3702)

Adekambi,T., Drancourt,M. and Raoult,D.

Submitted (26-MAR-2003) Unite des Rickettsies, Faculte de Medecine,

27 Boulevard Jean Moulin, Marseille 13385, France

Location/Qualifiers

1..3702

/organism="Mycobacterium wolinskyi"

/mol\_type="genomic DNA"

/strain="ATCC 700010"

/db\_xref="ATCC:700010"

/db\_xref="taxon:59750"

/note="type strain of Mycobacterium wolinskyi"

193..3702

/codon\_start=1

/transl\_table=1

/product="RpoB"

/protein\_id="AA076602.1"

/db\_xref="GI:34595759"

/translation="MLEGGLAVSSGSSNAITNNVGPANRVPAKLREPLFVPLGL  
LDVOTDSFEMLVGSGRQALIDGRENPGGLEEVLAELSPIIDFSGMSLSDDPR  
FDEYKASVDECKDKMTYAPLFTVTAELINNTEISIKSVTFMGDPFWMTKEGFTIN  
RGRQPVTVLLKALGNTNEQIVERRFSGEIMMGKIDKTSIGTDEALDITRKLRPG  
PTKESQTLLENLFFKEKRYDLARVGRYKINKKLGNAGKPIETEGCNIGLIGSL  
YLRLHGGOTMTVPGVEVVDIDHFGNRRRLRTVGLIQLVIRVGLSMEVREEMT  
DDVATIPOTLINRPPVAAIKIEFPGTSQLSQFOMDONNLSGLTHRRRLSALG  
PGSLSERGLVGRVYKINKKLGIATGLPIEASTLTEDDITVITTEYLRVLRAGTMMT  
APGVEPVVVDIDHFGNRRRLRTVGLIQLVIRVGLSMEVREEMTDDVATIPOT  
KVDGVVTDIIDLDEEDRRHVAAQANSPYDANGRTERTERVLRKRGSEVERFSSD

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3702;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCG 16

Db 1108 TGTAGCGGCCACGCG 1093

## RESULT 13

AY147167/c

LOCUS 3868 bp DNA linear BCT 01-SEP-2003  
DEFINITION Mycobacterium septicum strain ATCC 700731T RpoB (rpoB) gene,  
complete cds.

AY147167

VERSION

AY147167.1 GI:34391552

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

2 (bases 1 to 3868)

Adekambi,T., Drancourt,M. and Raoult,D.

Submitted (03-SEP-2002) Unite des Rickettsies, Faculte de Medecine,

Boulevard Jean Moulin, Marseille 13285, France

Location/Qualifiers

1..3868

/organism="Mycobacterium septicum"

/mol\_type="genomic DNA"

/strain="ATCC 700731T"

/db\_xref="ATCC:700731T"

/db\_xref="taxon:98668"

/gene="rpoB"

118..3675

/gene="rpoB"

118..3675

/codon\_start=1

/transl\_table=1

/product="RpoB"

/protein\_id="AA065354.1"

/db\_xref="GI:34391553"

/translation="MAANGPAVQLSGKSHVEEGCLAVSSGOSTANANTNNVGPAN  
RVSEFAKLREPLFVPLGLDVTDSFDMVAGDGRQAVDGRGEPNPGGLEEVLELSP  
IEDFSGMSLSDDPRFDEYKASVDECKDKMTYAPLFTVTAELINNTEISIKSVTF  
MGDPFWMTKEGFTINRGRQPVTVLLKALGNTNEQIVERRFSGEIMMGKIDKTSIG  
TDEALDITRKLRPGPTKESQTLLENLFFKEKRYDLARVGRYKINKKLGNAGKPIET  
EGCNIGLIGSLYLRLHGGOTMTVPGVEVVDIDHFGNRRRLRTVGLIQLVIRVGLS  
MEVREEMTDDVATIPOTLINRPPVAAIKIEFPGTSQLSQFOMDONNLSGLTHRRRL  
SALGPGSLSERGLVGRVYKINKKLGIATGLPIEASTLTEDDITVITTEYLRVLRAG  
TMMTAPGVEPVVVDIDHFGNRRRLRTVGLIQLVIRVGLSMEVREEMTDDVATIPOT  
KVDGVVTDIIDLDEEDRRHVAAQANSPYDANGRTERTERVLRKRGSEVERFSSD

ATLNSRLVEBDVLSIHEHEIDARDTKAGEETTRDIPNVSDVADLBERGIVR  
IAEVRDGDILVKGTPKGETELTPEERLRLAIFGKAEVRDTSIAKPHGSGKATIG  
IIVFSEDEDDDELPAVNEELVRYVVAQKRI SDGDKLAGRNGVIGKLLPVDMFPM  
PGTVPDIIINTHGVPRRNMIQOILETHLGMVAKAMINDVAGETPMASLPDGLHS  
APVDISVTPVPGDAGEELAGLSTLPPRODDVMVNAADGKATLFGEMSGEFPYPA  
TNGYMIILKHLHYDKDIKHARSNGPYSMITGOPLGGAGCGORFGEMCWMAOAAGA  
ATYLOELITIKSDTIVGRVYVAIKENIPEPBGIPESKUTLKLQSLCLANVEVLS  
SDGAALEMBDGDEDLERAANIGINLSRNSASVEDLA"

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3868;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16  
DB 1081 TGTAGCGGCCCGCG 1066

RESULT 14  
AY147173/c 3906 bp DNA linear BCT 01-SEP-2003  
LOCUS Mycobacterium fortuitum strain ATCC 49403T RpoB (rpoB) gene,  
DEFINITION complete cde.  
ACCESSION AY147173  
VERSION AY147173.1 GI:34391564  
KEYWORDS  
SOURCE  
ORGANISM Mycobacterium fortuitum  
Mycobacterium fortuitum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 3906)  
Adekambi, T., Drancourt, M. and Raoult, D.  
Taxonomy and identification of Mycobacteria from Mycobacterium  
fortuitum complex by complete sequencing of the rpoB gene  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 3906)  
Adekambi, T., Drancourt, M. and Raoult, D.  
REFERENCE Adekambi, T., Drancourt, M. and Raoult, D.  
AUTHORS Direct Submission  
TITLE Submitted (03-SEP-2002) Unite des Rickettsies, Faculte de Medecine,  
JOURNAL Boulevard Jean Moulin, Marseille 13385, France  
FEATURES  
source  
Location/Qualifiers  
1. .3906  
/organism="Mycobacterium fortuitum"  
/mol\_type="genomic DNA"  
/strain="ATCC 49403T"  
/db\_xref="ATCC:49403T"  
/db\_xref="taxon:1766"  
170. .3703  
/gene="rpoB"  
170. .3703  
/gene="rpoB"  
/codon\_start=1  
/transl\_table=1  
/product="RpoB"  
/protein\_id="AA05360.1"  
/db\_xref="GI:34391565"

/transl\_table="MSEKSHVLEGCILAVSSOSKSNATITNNSVPGAENRVSFAKL  
EPLPGLADVOTSPFVLVSDERQAVDRGEPDGLAEVLELSELPEDSGSM  
SLSFDPDFDEKAVVDECKDMYTAAPLVTAFFINNNGEIKSQTVPMGDPFMMT  
EKGTIFINGRQVVSQVLRSPGVFDESIDSTEKTLHSKVIPGRAMLEFVDKRA  
DTGVARIDRKROPTVLLKALGMTEQITERFSGSEIMGTLESDTAGDEALDI  
YKLRGPBPPTKESQOTLENIFFEKERYDLARVGRVYKMKLGLAAGQPTSTLTL  
EDVATTEVLRHGGOTTMTVPGVEVPEVDDMDHGRRLRTVGBLIONOIRVGL  
SRMERVAREMTTQDVEAITPOTLININPVAIAITEPCTSOLOPMONNIGSLTH  
RRRLSALDPGLSREKGLVARDVSHVGMCPITEPGENIGLISLSTYARINPE  
GFIETPRYKVVGVVTDQIDYLTADBEDRHAQAANSPIIDGARTBRLVWVRSGS  
VENVAISPVDVMDVSPROMVATAMIPLFLEHDDNRAIMGMANMORQAVPLVRSAPL  
VGTGMEELAAIDAGVIVSEKAGIIEVSADYITVADGSHITRYMKFARSNGTC  
ANORPIVAGORVEAGOVVADGCTENGEMALGKULVAIMPMBEGNEDAIILSNRL  
VEBDVLSIHEHEIDARDTKAGEETTRDIPNVSDVADLBERGIVRGAERD  
DILVKGTPKGETELTPEERLRLAIFGKAEVRDTSIAKPHGSGKATIGRVSRED  
DDELPAVNEELVRYVVAQKRI SDGDKLAGRNGVIGKLLPVDMFPMGTPVDI

IINTHGVPRRNMIQOILETHLGMVAKAMINDVAGETPMASLPOLYSAPVDSIVS  
TPVPGDAGEELAGLSTLPPRODDVMVNAADGKATLFGEMSGEFPYPAVYMYTL  
KLHLHYDKDIKHARSNGPYSMITGOPLGGAGCGORFGEMCWMAOAAYATLOELI  
TIKSDTIVGRVYVAIKENIPEPBGIPESKUTLKLQSLCLANVEVLSDDGAALIE  
RQDDEDLERAANIGINLSRNSASVEDLA"

## ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3906;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16  
DB 1109 TGTAGCGGCCCGCG 1094

RESULT 15  
AY262736/c 3929 bp DNA linear BCT 13-JAN-2004  
LOCUS Mycobacterium goodii strain ATCC 700504 RpoB gene, complete cde.  
DEFINITION AY262736  
ACCESSION AY262736  
VERSION AY262736.1 GI:34595744  
KEYWORDS  
SOURCE  
ORGANISM Mycobacterium goodii  
Mycobacterium goodii  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 3929)  
Adekambi, T., Colson, P. and Drancourt, M.  
rpoB-based identification of nonpigmented and late-pigmenting  
rapidly growing mycobacteria  
JOURNAL J. Clin. Microbiol. 41 (12), 5659-5708 (2003)  
AUTHORS 2 (bases 1 to 3929)  
Adekambi, T., Drancourt, M. and Raoult, D.  
REFERENCE Adekambi, T., Drancourt, M. and Raoult, D.  
AUTHORS Direct Submission  
TITLE Submitted (26-MAR-2003) Unite des Rickettsies, Faculte de Medecine,  
JOURNAL 27 Boulevard Jean Moulin, Marseille 13385, France  
FEATURES  
source  
Location/Qualifiers  
1. .3929  
/organism="Mycobacterium goodii"  
/mol\_type="genomic DNA"  
/strain="ATCC 700504"  
/db\_xref="ATCC:700504"  
/db\_xref="taxon:134601"  
/note="type strain of Mycobacterium goodii"  
193. .3702  
/codon\_start=1  
/transl\_table=1  
/product="RpoB"  
/protein\_id="AA076595.1"  
/db\_xref="GI:34595745"

/transl\_table="MLEGCILAVSSOSKSNATITNNSVPGAENRVSFAKLREPLEVGL  
LDVOTSPFVLVSDERQAVDRGEPDGLAEVLELSELPEDSGSMISLSDPR  
PDEYKASVQVDECKDMYTAAPLVTAFFINNNGEIKSQTVPMGDPFMMT  
EKGTIFINGRQVVSQVLRSPGVFDESIDSTEKTLHSKVIPGRAMLEFVDKRA  
DTGVARIDRKROPTVLLKALGMTEQITERFSGSEIMGTLESDTAGDEALDI  
YKLRGPBPPTKESQOTLENIFFEKERYDLARVGRVYKMKLGLAAGQPTSTLTL  
EDVATTEVLRHGGOTTMTVPGVEVPEVDDMDHGRRLRTVGBLIONOIRVGL  
SRMERVAREMTTQDVEAITPOTLININPVAIAITEPCTSOLOPMONNIGSLTH  
RRRLSALDPGLSREKGLVARDVSHVGMCPITEPGENIGLISLSTYARINPE  
GFIETPRYKVVGVVTDQIDYLTADBEDRHAQAANSPIIDGARTBRLVWVRSGS  
VENVAISPVDVMDVSPROMVATAMIPLFLEHDDNRAIMGMANMORQAVPLVRSAPL  
VGTGMEELAAIDAGVIVSEKAGIIEVSADYITVADGSHITRYMKFARSNGTC  
ANORPIVAGORVEAGOVVADGCTENGEMALGKULVAIMPMBEGNEDAIILSNRL  
VEBDVLSIHEHEIDARDTKAGEETTRDIPNVSDVADLBERGIVRGAERD  
DILVKGTPKGETELTPEERLRLAIFGKAEVRDTSIAKPHGSGKATIGRVSRED  
DDELPAVNEELVRYVVAQKRI SDGDKLAGRNGVIGKLLPVDMFPMGTPVDI

ERRAANIGINLSRNSASVEDLA"

ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 3929;

Best Local Similarity 100.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGGCCACGCG 16

Db 1108 TGTAGCGGGCCACGCG 1093

Search completed: February 10, 2005, 13:13:45  
Job time : 120.196 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 19.6391 Seconds  
(without alignments)  
5727.097 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19

Sequence: 1 tctagcgccgcccagcgctg 19

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 100.0 | 19          | 6      | ABK15246  | ABK15246 Corn tran  |
| 2          | 100.0 | 498         | 6      | ABK15244  | ABK15244 Corn nk60  |
| 3          | 84.2  | 208         | 4      | AAS05219  | AAS05219 Mycobacte  |
| 4          | 84.2  | 35167       | 12     | ADMA43175 | ADMA43175 Wild-type |
| 5          | 84.2  | 35167       | 13     | ADS00140  | ADS00140 Wild-type  |
| 6          | 78.9  | 752         | 4      | AA159900  | AA159900 Human pol  |
| 7          | 78.9  | 1098        | 12     | AD030319  | AD030319 Mouse GPC  |
| 8          | 78.9  | 1183        | 12     | ADP72493  | ADP72493 Renal tox  |
| 9          | 78.9  | 1253        | 10     | ADBS2577  | ADBS2577 Primary r  |
| 10         | 78.9  | 1405        | 2      | AAQ46125  | AAQ46125 PG82 rece  |
| 11         | 78.9  | 1780        | 10     | ABV75862  | ABV75862 Human pot  |
| 12         | 78.9  | 2107        | 2      | AAQ46124  | AAQ46124 PG82 rece  |
| 13         | 78.9  | 14800       | 10     | ABV75863  | ABV75863 Human pot  |
| 14         | 78.9  | 51855       | 10     | ACF04818  | ACF04818 Melithiaz  |
| 15         | 78.9  | 214520      | 10     | ADL13471  | ADL13471 Osecoarh   |
| 16         | 73.7  | 100         | 8      | ACD72269  | ACD72269 E. coli K  |
| 17         | 73.7  | 228         | 8      | ABX54979  | ABX54979 Bovine ES  |
| 18         | 73.7  | 228         | 13     | ADBS5656  | ADBS5656 Bacterial  |
| 19         | 73.7  | 292         | 8      | ABX42572  | ABX42572 Bovine ES  |
| 20         | 73.7  | 414         | 9      | ACH29799  | ACH29799 Human tes  |

# ALIGNMENTS

|    |    |      |      |    |           |                     |
|----|----|------|------|----|-----------|---------------------|
| 21 | 14 | 73.7 | 436  | 8  | ABX52378  | ABX52378 Bovine ES  |
| 22 | 14 | 73.7 | 465  | 12 | ACH92095  | ACH92095 Human gen  |
| 23 | 14 | 73.7 | 500  | 4  | AAD10051  | AAD10051 Mouse 5'   |
| 24 | 14 | 73.7 | 619  | 13 | ADR65184  | ADR65184 Cotton CD  |
| 25 | 14 | 73.7 | 756  | 12 | ADK13721  | ADK13721 E. coli i  |
| 26 | 14 | 73.7 | 834  | 12 | ADJ39824  | ADJ39824 Plant cDN  |
| 27 | 14 | 73.7 | 862  | 5  | AAS74970  | AAS74970 DNA encod  |
| 28 | 14 | 73.7 | 870  | 12 | ADJ43711  | ADJ43711 Plant cDN  |
| 29 | 14 | 73.7 | 1004 | 5  | AAS74969  | AAS74969 DNA encod  |
| 30 | 14 | 73.7 | 1146 | 8  | ACA26804  | ACA26804 Prokaryot  |
| 31 | 14 | 73.7 | 1209 | 8  | AB198018  | AB198018 Non-endog  |
| 32 | 14 | 73.7 | 1209 | 8  | ABV76938  | ABV76938 Nucleotid  |
| 33 | 14 | 73.7 | 1209 | 12 | ADO30027  | ADO30027 Human GPC  |
| 34 | 14 | 73.7 | 1226 | 3  | AA293894  | AA293894 Human EP-  |
| 35 | 14 | 73.7 | 1376 | 3  | AAA34920  | AAA34920 Human ade  |
| 36 | 14 | 73.7 | 1376 | 3  | AA221042  | AA221042 Human low  |
| 37 | 14 | 73.7 | 1376 | 8  | AB234821  | AB234821 Coding se  |
| 38 | 14 | 73.7 | 1376 | 8  | AB242749  | AB242749 Human pro  |
| 39 | 14 | 73.7 | 1376 | 10 | ABA296736 | ABA296736 Human nuc |
| 40 | 14 | 73.7 | 1376 | 13 | ACA56763  | ACA56763 Human sig  |
| 41 | 14 | 73.7 | 1376 | 11 | ABD20585  | ABD20585 Human pul  |
| 42 | 14 | 73.7 | 1376 | 12 | AD156559  | AD156559 Human pol  |
| 43 | 14 | 73.7 | 1394 | 13 | ADR46583  | ADR46583 Human pro  |
| 44 | 14 | 73.7 | 1394 | 2  | AAQ80287  | AAQ80287 Prostagla  |
| 45 | 14 | 73.7 | 1431 | 8  | ACA23808  | ACA23808 Prokaryot  |

# RESULT 1

ABK15246 standard; DNA; 19 BP.

ABK15246;

08-MAY-2002 (first entry)

Corn transgene junction probe #1.

Probe; ss; rice actin 1 promoter; RAI; RAI intron;  
KM chloroplast transit peptide gene; glyphosate resistance; corn;  
KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
KW transcripional terminator; cauliflower mosaic virus 35S promoter;  
KM PV-ZWGT32; transgenic; nk603.

Zea mays.  
OS Synthetic.

PN EPI167531-A1.

PD 02-JAN-2002.

PE 15-JUN-2001; 2001EP-00202314.

PR 22-JUN-2000; 2000US-0213567P.

PR 13-OCT-2000; 2000US-0240014P.

PR 13-OCT-2000; 2000US-0241215P.

PA (MONS ) MONSANTO TECHNOLOGY LLC.

PI Behr CF, Hironaka C, Heck GR, You J;

PT WPI; 2002-165871/22.

PT Novel DNA construct useful for producing a corn plant that tolerates application of glyphosate herbicide, comprises two transgene expression cassettes.

PS Claim 8; Page 3; 25pp; English.

CC The invention relates to a DNA construct (I) comprising a first (F1) and a second (F2) expression cassette. F1 of the DNA construct, in operable

CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
 CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
 CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
 CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
 CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
 CC PV-ZMG732(nk603) of 498 or 1183bp, or primers and probes derived from  
 CC them (used to detect transgene junction points). The construct is useful  
 CC for producing a corn plant that tolerates application of glyphosate  
 CC herbicide by transforming a corn cell with, selecting the corn cell for  
 CC tolerance to application of glyphosate, and growing the corn cell into a  
 CC fertile corn plant. The present sequence is a transgene junction probe  
 CC sequence consisting of corn genomic DNA and vector sequences  
 CC  
 SQ Sequence 19 BP; 2 A; 6 C; 8 G; 3 T; 0 U; 0 Other;  
 XX  
 Query Match 100.0%; Score 19; DB 6; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.14; Mismatches 0; Gaps 0;  
 Matches 19; Conservative 0; Indels 0; Indels 0; Gaps 0;  
 OY 1 TGTAGCGGCCCGCGGTGG 19  
 1 TGTAGCGGCCCGCGGTGG 19  
 DB 1 TGTAGCGGCCCGCGGTGG 19  
 RESULT 2  
 ID ABLK5244 standard; DNA; 498 BP.  
 AC ABLK5244;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Corn nk603 5' transgene/genomic sequence PCR fragment.  
 XX  
 KW db; rice actin 1 promoter; RA1; RA1 intron;  
 KW chloroplast transit peptide gene; glyphosate resistance; corn;  
 KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
 KW transcriptional terminator; cauliflower mosaic virus 35S promoter;  
 KW PV-ZMG732; transgenic; nk603.  
 XX  
 OS Zea mays.  
 OS Oryza sativa.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 FT Key Location/Qualifiers  
 FT misc\_feature 1..304  
 FT /\*tag= a  
 FT /note= "Corn genomic sequence"  
 FT misc\_feature 305..349  
 FT /\*tag= b  
 FT /note= "Vector sequence"  
 FT misc\_feature 350..498  
 FT /\*tag= c  
 FT /note= "Rice ACT1 sequence"  
 XX  
 PN EP1167531-A1.  
 PD 02-JAN-2002.  
 XX  
 PF 15-JUN-2001; 2001EP-00202314.  
 XX  
 PR 22-JUN-2000; 2000US-0213567P.  
 PR 13-OCT-2000; 2000US-0240014P.  
 PR 13-OCT-2000; 2000US-0241215P.  
 XX  
 PA (MONS ) MONSANTO TECHNOLOGY LLC.  
 XX  
 PI Behr CF, Hironaka C, Heck GR, You J;  
 XX  
 DR WPI, 2002-165871/22.  
 XX  
 PT Novel DNA construct useful for producing a corn plant that tolerates

PT application of glyphosate herbicide, comprises two transgene expression  
 PT cassettes.  
 XX  
 PS Claim 3; Page 16; 25pp; English.  
 XX  
 CC The invention relates to a DNA construct (I) comprising a first (F1) and  
 CC a second (F2) expression cassette. F1 of the DNA construct, in operable  
 CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
 CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
 CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
 CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
 CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
 CC PV-ZMG732(nk603) of 498 or 1183bp, or primers and probes derived from  
 CC them (used to detect transgene junction points). The construct is useful  
 CC for producing a corn plant that tolerates application of glyphosate  
 CC herbicide by transforming a corn cell with, selecting the corn cell for  
 CC tolerance to application of glyphosate, and growing the corn cell into a  
 CC fertile corn plant. The present sequence is the 498bp amplicon consisting  
 CC of corn genomic sequence, vector sequences and Rice ACT1 promoter  
 CC sequences  
 CC  
 SQ Sequence 498 BP; 148 A; 114 C; 131 G; 105 T; 0 U; 0 Other;  
 XX  
 Query Match 100.0%; Score 19; DB 6; Length 498;  
 Best Local Similarity 100.0%; Pred. No. 0.097;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 TGTAGCGGCCCGCGGTGG 19  
 296 TGTAGCGGCCCGCGGTGG 314  
 DB 296 TGTAGCGGCCCGCGGTGG 314  
 RESULT 3  
 ID AAS05219/c  
 AC AAS05219; standard; DNA; 208 BP.  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Mycobacterium fortuitum rpoB gene fragment.  
 XX  
 KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PPA; RFLP;  
 KW PCR-restriction fragment length polymorphism analysis; ds.  
 XX  
 OS Mycobacterium fortuitum.  
 XX  
 PN WO200131061-A1.  
 PD 03-MAY-2001.  
 XX  
 PR 27-OCT-2000; 2000WO-KR001223.  
 PR 27-OCT-1999; 99KR-00046795.  
 XX  
 PA (ERUM-) ERUME BIOTECH CO LTD.  
 XX  
 PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;  
 XX  
 DR WPI; 2001-300520/31.  
 XX  
 PT New DNA fragments from the rpoB gene of mycobacteria, useful for  
 PT diagnosis and identification of many mycobacterial species by restriction  
 PT fragment length polymorphism.  
 XX  
 PS Claim 1; Page 46; 50pp; English.  
 XX  
 CC The present sequence for Mycobacterium fortuitum rpoB gene fragment is 1  
 CC of 24 rpoB gene fragments (AAS05201-AAS05224) from various Mycobacterial  
 CC species. These rpoB gene fragments can be used in the diagnosis and  
 CC identification of Mycobacterium species using a novel PCR-restriction  
 CC fragment length polymorphism analysis (PRA) method. The method comprises  
 CC obtaining a restriction fragment length polymorphism (RFLP) pattern of

CC the 24 rpoB gene fragments, isolating, amplifying and digesting the DNA  
 CC fragment from the microorganism to be identified and comparing the RFLP  
 CC patterns from the known rpoB gene fragments with the unidentified  
 CC fragment. The rpoB gene fragments are useful to identify a wide range of  
 CC Mycobacterium species, e.g. for diagnosis or to obtain epidemiological  
 CC and pathogenesis information for selection of appropriate therapies  
 CC including M. tuberculosis, M. lepre and non-tuberculous mycobacteria  
 CC (NTM) encountered in subjects infected with human immunodeficiency virus  
 CC (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and  
 CC cost effective (only 1 PCR required), and can differentiate between many  
 CC species in a single experiment, including those difficult to distinguish  
 CC by usual biochemical tests. Also described are oligonucleotide probes  
 CC (AAS05227-AAS05242) for detecting specific Mycobacterial species

SQ Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 4; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 6.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16  
 DB 42 TGTAGCGGCCCGCG 27

RESULT 4

ID ADM43175/c  
 ID ADM43175 standard; DNA; 35167 BP.

AC ADM43175;

DT 03-JUN-2004 (first entry)

XX Wild-type adenovirus serotype 24 nucleotide sequence SEQ ID NO:5.

KM replication-defective adenovirus; adenoviral E1-complementing cell line;

KM anti-HIV; gene therapy; vaccine; immune response; HIV;

XX adenovirus serotype 24; gene; ds.

OS unidentified adenovirus.

XX WO2004018627-A2.

PD 04-MAR-2004.

PE 21-AUG-2003; 2003WO-US026145.

XX 22-AUG-2002; 2002US-0405182P.

PR 17-MAR-2003; 2003US-0455234P.

PR 17-MAR-2003; 2003US-0455312P.

PR 26-MAR-2003; 2003US-0458825P.

XX (MERI ) MERCK & CO INC.

PI Belt AJ, Chastain M, Sandig V, Emin EA, Shiver JW, Casimiro DR;

PI Kaelow DC, Morsey M;

XX WPI; 2004-226820/21.

DR Propagating replication-defective adenovirus in an adenoviral E1-

PT complementing cell line, useful in gene therapy applications, comprises

PT providing all or a portion of an E4 region in cis within the genome of

PT the adenovirus.

XX Example 7; SEQ ID NO 5; 113bp; English.

CC is propagated by the method described above; (2) a recombinant adenoviral  
 CC vector of serotype 24 which comprises an E4 gene or a segment of an E4  
 CC gene comprising ORF6 of an alternative serotype; (3) a population of  
 CC cells comprising the recombinant adenoviral vector; (4) producing  
 CC recombinant, replication-defective adenovirus particles; (5) a purified  
 CC recombinant, replication-defective adenovirus particles harvested in  
 CC accordance with the method in (4); (6) a composition comprising purified  
 CC recombinant adenovirus particles; (7) effecting the delivery and  
 CC expression of heterologous nucleic acid; and (8) generating a cellular-  
 CC mediated immune response against HIV in an individual. The replication-  
 CC defective adenovirus has anti-HIV activity, and can be used in gene  
 CC therapy and in vaccines. The method can be used for propagating a  
 CC replication-defective adenovirus in an adenoviral E1-complementing cell  
 CC line useful for propagating and rescuing multiple serotypes of  
 CC replication-defective adenovirus in a single adenoviral E1-complementing  
 CC cell line and for generating a cellular-mediated immune response against  
 CC HIV in an individual. The resultant virus can be studied and used in  
 CC various gene therapy and vaccine efforts. The present sequence represents  
 CC a wild-type adenovirus serotype 24 nucleotide sequence, which is used in  
 CC the exemplification of the present invention.

SQ Sequence 35167 BP; 7977 A; 9958 C; 9996 G; 7236 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 12; Length 35167;  
 Best Local Similarity 100.0%; Pred. No. 3.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGCG 16  
 DB 12091 TGTAGCGGCCCGCG 12076

RESULT 5

ID ADS00140/c  
 ID ADS00140 standard; DNA; 35167 BP.

AC ADS00140;

DT 16-DEC-2004 (first entry)

XX Wild-type human adenovirus serotype 24 DNA SEQ ID NO:1.

KM de; adenovirus serotype 24; anti-HIV; vector; gene therapy; vaccine.

XX Human adenovirus type 24.

XX WO2004083418-A1.

PD 30-SEP-2004.

PE 21-AUG-2003; 2003WO-US026338.

PR 17-MAR-2003; 2003US-0455312P.

XX (MERI ) MERCK & CO INC.

PI Emin EA, Shiver JW, Belt AJ, Casimiro DR, Chastain M, Kaelow DC;

PI Morsey MA;

XX WPI; 2004-691052/67.

DR New recombinant adenoviral vector of serotype 24 which is partially

PT deleted in E1 and devoid of E1 activity, useful for delivering and

PT expressing exogenous genes, e.g. in gene therapy or in vaccination

PT protocols.

XX Disclosure; SEQ ID NO 1; 133bp; English.

XX The invention relates to a novel recombinant adenoviral vector of

XX serotype 24 which is partially deleted in E1 and devoid of E1 activity

XX and comprises a heterologous nucleic acid or an human immunodeficiency

XX virus (HIV)-1 gene. A vector of the invention has anti-HIV activity, and

XX may have a use in gene therapy, and as a vaccine. The recombinant

CC adenoviral vector of serotype 24 which is at least partially deleted in  
CC E1 and devoid of E1 activity is useful for the delivery and expression of  
CC exogenous genes. The adenoviral vectors are useful in gene therapy or in  
CC vaccination protocols. The present sequence represents the nucleic acid  
CC sequence for wild-type adenovirus serotype 24.

XX SQ Sequence 35167 BP; 7977 A; 9958 C; 9996 G; 7236 T; 0 U; 0 Other;

Query Match 84.2%; Score 16; DB 13; Length 35167;  
Best Local Similarity 100.0%; Pred. No. 3.4;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGCCCGCCG 16

DB 12091 TGTAGCGCCCGCCG 12076

RESULT 6  
AA159900/c  
ID AA159900 standard; cDNA; 752 BP.

AC AA159900;

DT 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 3889.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX peripheral nervous system; neuropathy; central nervous system; CNS;  
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;  
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;  
XX leukaemia; ss.

OS Homo sapiens.

PN WO20015312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-0052317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 29-NOV-2000; 2000US-00727344.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
XX P-PSDB; AAM40744.

PT Novel nucleic acids and polypeptides, useful for treating disorders such  
as central nervous system injuries.

PS Claim 1; SEQ ID NO 3889; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
CC encoded polypeptides (AAM38642-AA42213) with noctropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemoreactive/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification

XX SQ Sequence 752 BP; 184 A; 156 C; 233 G; 179 T; 0 U; 0 Other;

Query Match 78.9%; Score 15; DB 4; Length 752;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGCCCGCCGCT 17

DB 30 TAGCGCCCGCCGCT 16

RESULT 7  
AD030319/c  
ID AD030319 standard; cDNA; 1098 BP.

AC AD030319;

DT 29-JUL-2004 (first entry)

DE Mouse GPCR PIGR3 polynucleotide, SEQ ID NO:1422.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;  
XX transgenic mouse; neurological disorder; adrenal gland disorder;  
XX colon disorder; intestinal disorder; cardiovascular disorder;  
XX muscular disorder; blood disorder; immune disorder; bone disorder;  
XX joint disorder; metabolic disorder; nutritive disorder; cancer;  
XX kidney disorder; liver disorder; lung disorder; breast disorder;  
XX ovary disorder; uterus disorder; prostate disorder; testis disorder;  
XX skin disorder; stomach disorder; pancreas disorder; spleen disorder;  
XX thymus disorder; thyroid disorder; antiparkinsonian; antianemic;  
XX cyostatic; antiinflammatory; vasotropic; antiangiogenic; antidiabetic;  
XX CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;  
XX vitruclide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;  
XX dermatological; antitumor; antithyroid; antiallergic; anorectic;  
XX immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;  
XX murine; gene; ss.

OS Mus musculus.

PN WO2004040000-A2.

XX 13-MAY-2004.

PF 09-SEP-2003; 2003WO-US028226.

XX 09-SEP-2002; 2002US-0409303P.  
XX 09-APR-2003; 2003US-0461329P.

PA (PRIM-) PRIMAL INC.

XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;  
PI Madison L, McIlwain KL, Pavlova KM, Vasiliadis D, Zeng H;  
XX  
XX WPI: 2004-390329/36.  
XX P-PSDB; ADO29621.

PT Novel mammalian G protein coupled receptors, useful for identifying  
PT compounds that modulates diagnosing and treating disease condition  
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina  
PT pectoris, Parkinson's disease.

PS Claim 151; SEQ ID NO 1422; 542bp; English.

XX The invention relates to human and mouse G protein-coupled receptors



(GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several mice, each of which has a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridise to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia); disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and thyroid (e.g., cancers). The present sequence represents a GPCR-encoding nucleic acid of the invention. Note: The full sequence data for this patent did not form part of the printed specification; those sequences not shown were obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 1098 BP; 185 A; 342 C; 335 G; 236 T; 0 U; 0 Other;  
Query Match 78.9%; Score 15; DB 12; Length 1098;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGC 15  
Db 526 TGTAGCGGCCACGC 512

RESULT 8  
ADP72493/c  
ID ADP72493 standard; DNA; 1183 BP.

AC ADP72493;

DT 26-AUG-2004 (first entry)

DE Renal toxin progression gene marker #1082.

XX db; toxic effect; gene expression profile; kidney tissue;  
KM differential gene expression; toxicity progression; toxicity marker;  
KM drug screening; toxicity assay; kidney pathology; nephritis;  
KM kidney necrosis; glomerular injury; tubular injury;  
KM focal segmental glomerulosclerosis.

XX Rattus norvegicus.

OS WO2004048598-A2.

PN 10-JUN-2004.

PD 24-NOV-2003; 2003WO-US037556.

PF 22-NOV-2002; 2002US-00301856.

PR (GENE-) GENE LOGIC INC.

XX Mendrick DL, Porter MM, Johnson KR, Castle A, Higgs B;  
PI Elsholtz M;

DR WPI; 2004-460771/43.

XX predicting (the progression of) a toxic effect of a compound, for  
PT monitoring the progression of renal disease states, comprises preparing a  
PT gene expression profile of a kidney tissue or cell sample exposed to the  
PT compound.

PS Claim 11; SEQ ID NO 1082; 266pp; English.

XX The invention relates to a method of predicting (the progression of) a  
XX toxic effect of a compound by preparing a gene expression profile of a  
XX kidney tissue or cell sample exposed to the compound and comparing the  
XX gene expression profile to a database, or detecting the level of gene(s)  
XX expression in a tissue or cell sample exposed to the compound, where  
XX differential gene expression compared to a control indicates a toxic  
XX effect (toxicity progression). The method is useful for predicting (the  
XX progression of) at least one toxic effect of a compound. The genes are  
XX useful as toxicity markers in drug screening and toxicity assays. The  
XX methods are useful for predicting the likelihood that a compound or test  
XX agent will induce various specific kidney pathologies, such as nephritis,  
XX kidney necrosis, glomerular and tubular injury, or focal segmental  
XX glomerulosclerosis. The methods are useful for determining the similarity  
XX of a toxic response to one or more individual compounds and for  
XX predicting or elucidating the potential cellular pathways influenced,  
XX induced or modulated by the compound or test agent. The kit is useful for  
XX predicting or modelling the toxic response of a test compound, for  
XX monitoring the progression of renal disease states, for identifying genes  
XX that show promise as new drug targets and for screening known and newly  
XX designed drugs. This sequence corresponds to a gene marker used in the  
XX method of the invention. (Note: The sequence data for this patent did not  
XX form part of the printed specification, but was obtained in electronic  
XX format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences).

Sequence 1183 BP; 208 A; 361 C; 366 G; 248 T; 0 U; 0 Other;  
Query Match 78.9%; Score 15; DB 12; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGC 15  
Db 582 TGTAGCGGCCACGC 568

RESULT 9  
ADB52577/c  
ID ADB52577 standard; DNA; 1253 BP.

AC ADB52577;

DT 04-DEC-2003 (first entry)

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3119.

XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;  
KM toxicity marker; toxicity progression; drug screening;  
KM primary rat hepatocyte toxicity modelling; gene; ds.

XX Rattus norvegicus.

OS WO2003065993-A2.

PN 14-AUG-2003.

PD 04-FEB-2003; 2003WO-US003482.

PF 04-FEB-2002; 2002US-035317P.

PR 13-MAR-2002; 2002US-036353P.

PR 08-APR-2002; 2002US-037024P.

PR 10-APR-2002; 2002US-0371134P.

PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.

PR 11-APR-2002; 2002US-0371413P.

```
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castele A, Orr M;
XX Elashoff M;
XX WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
XX identifying hepatotoxic compounds, comprises comparing a gene expression
XX profile of a tissue or cell sample to a database of tox mean and non-tox
XX mean values.
XX
XX Claim 44; SEQ ID NO 3119; 874bp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the tox mean and non-tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity
XX markers for the prediction or identification of the physiological state
XX of tissue or cell sample that has been exposed to a compound, or to
XX identify or predict the toxic effects of a compound or an agent. These
XX CC may also be used as markers for monitoring toxicity progression or for
XX drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
XX
XX SQ Sequence 1253 BP; 213 A; 386 C; 373 G; 281 T; 0 U; 0 Other;
XX
XX Query Match 78.9%; Score 15; DB 10; Length 1253;
XX Best Local Similarity 100.0%; Pred. No. 19;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TGTAGCGGCCGACGC 15
XX |||||
XX 614 TGTAGCGGCCGACGC 600
XX
XX Db
XX
XX RESULT 10
XX AA046125/c
XX ID AA046125 standard; cDNA; 1405 BP.
XX
XX AC AA046125;
XX
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 09-FEB-1994 (first entry)
XX
XX DE PGE2 receptor (EP3 beta) clone MP653.
XX
XX XX
XX KM prostaglandin B receptor; PGE; agonist; antagonist; GTP-binding protein;
XX KM digestive tract; constriction; relaxation; gastric acid;
XX KM intestinal juice; neurotransmitter; ss.
XX
XX OS Mus musculus.
XX
XX XX
XX Key Location/Qualifiers
XX FH 106..1191
XX FT /*tag= a
XX FT /product= "PGE2_receptor"
```

```
FT
XX
XX EP57966-A1.
XX
XX PD 01-SEP-1993.
XX
XX PF 24-FEB-1993; 93EP-00102873.
XX
XX PR 24-FEB-1992; 92JP-00036580.
XX PR 23-MAR-1992; 92JP-00064889.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX XX
XX PI Ichikawa A, Narumiya S;
XX
XX DR WPI; 1993-274435/35.
XX DR P-P5DB; AAR42281.
XX
XX XX
XX PT New prostaglandin E receptor protein and DNA encoding it - used to study
XX PT prostaglandin binding, agonists and antagonists etc.
XX
XX PS Claim 15; Fig 4; 22bp; English.
XX
XX PGE2 is known to be involved in digestive tract constriction/relaxation,
XX gastric juice secretion and neurotransmitter release. The protein encoded
XX by the gene is capable of receiving PGE and of activating GTP-binding
XX protein. MP653 (AA046125) is identical to MP660 (AA046124) with the
XX exception of a 89-bp deletion in the coding region of the C-terminal tail
XX of the receptor. This deletion creates a new reading frame downstream
XX from this junction. This results in a 30 amino acid C-terminal fragment
XX of the receptor being replaced with a new 26-amino acid fragment in the C
XX -terminal end of the MP653 (AA046125) receptor. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX SQ Sequence 1405 BP; 251 A; 415 C; 418 G; 321 T; 0 U; 0 Other;
XX
XX Query Match 78.9%; Score 15; DB 2; Length 1405;
XX Best Local Similarity 100.0%; Pred. No. 19;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TGTAGCGGCCGACGC 15
XX |||||
XX 631 TGTAGCGGCCGACGC 617
XX
XX Db
XX
XX RESULT 11
XX ABV75862/c
XX ID ABV75862 standard; cDNA; 1780 BP.
XX
XX AC ABV75862;
XX
XX DT 05-FEB-2003 (first entry)
XX
XX DE Human potassium channel transporter cDNA.
XX
XX KM Transporter; potassium channel; human; gene therapy; chromosome 17; gene;
XX KM ss.
XX
XX OS Homo sapiens.
XX
XX XX
XX Key Location/Qualifiers
XX FH 1..374
XX FT /*tag= a
XX FT 375..1667
XX FT CDS /*tag= b
XX FT /product= "Transporter protein"
XX FT 1667..1780
XX FT /*tag= c
XX
XX PD WO200281657-A2.
XX PD 17-OCT-2002.
```

PF 01-APR-2002; 2002WO-US009743.  
 XX 30-MAR-2001; 2001US-00820923.  
 XX (BEKE) PE CORP NY.  
 PA (BEAS/) BEASLEY E M.  
 PA (WEIM/) WEI M.  
 PA (KETC/) KETCHUM K A.  
 PA (DFRA/) DI FRANCESCO V.  
 XX  
 PI Beasley EM, Wei M, Ketchum KA, Di Francesco V;  
 XX  
 DR WPI; 2003-058533/05.  
 DR P-PSDB; ABP55112.  
 XX  
 PT Novel isolated human transporter peptide useful for treating disorder  
 PT characterized by absence of, in appropriate or unwanted expression of the  
 PT transporter protein, and as immunogens to raise antibodies.  
 XX  
 PS Claim 4; Fig 1; 75pp; English.  
 XX  
 CC The present sequence is that of cDNA encoding a novel human transporter  
 CC protein that is related to the voltage gated potassium channel  
 CC transporter family, and which is expressed in the lung and in the foetal  
 CC brain. The gene (see ABV75863) encoding the transporter is located on  
 CC chromosome 17. Nucleic acids encoding the transporter are useful as  
 CC probes, primers, in biological assays, for constructing recombinant  
 CC vectors, host cells and transgenic animals, for expressing antigenic  
 CC portions of the human transporter protein, for designing ribozymes, for  
 CC monitoring the effectiveness of modulating compounds on the expression or  
 CC activity of the human transporter gene in clinical trials or in a  
 CC treatment regimen, in diagnostic assays for qualitative changes in human  
 CC transporter nucleic acid that lead to pathology, for testing an  
 CC individual for a genotype that while not necessarily causing a disease,  
 CC nevertheless affects the treatment modality, as antisense constructs to  
 CC control human transporter gene expression in cells, tissues and  
 CC organisms, and for gene therapy in patients containing cells that are  
 CC aberrant in human transporter gene expression. A gene chip comprising a  
 CC transporter nucleic acid is claimed, and is useful for conducting cell-  
 CC based assays involving the human transporter protein  
 XX  
 XX Sequence 1780 BP; 352 A; 570 C; 496 G; 362 T; 0 U; 0 Other;  
 SQ  
 Query Match 78.9%; Score 15; DB 10; Length 1780;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 AGCGGCCCGCGGTG 18  
 Db 51 AGCGGCCCGCGGTG 37  
 RESULT 12  
 AAQ46124/c  
 ID AAQ46124 standard; cDNA; 2107 BP.  
 XX  
 AC AAQ46124;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 09-FEB-1994 (first entry)  
 XX  
 DE PGE2 receptor (EP3 alpha) clone MP660.  
 XX  
 KW prostaglandin E receptor; PGE; agonist; antagonist; GTP-binding protein;  
 KW digestive tract; constriction; relaxation; gastric acid;  
 KW intestinal juice; neurotransmitter; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 119..1216  
 FT /tag= a  
 FT /product= "PGE2\_receptor"

FT /note= "PGE subtype EP3 alpha"  
 XX  
 XX EP57966-A1.  
 XX  
 XX 01-SEP-1993.  
 PD  
 XX  
 XX 24-FEB-1993; 93EP-00102873.  
 PP  
 XX  
 XX 24-FEB-1992; 92UP-00036580.  
 PR 23-MAR-1992; 92UP-00064889.  
 XX  
 XX (TAKA ) TAKEDA CHEM IND LTD.  
 PA  
 XX  
 PI Ichikawa A, Naramiya S;  
 XX  
 XX WPI; 1993-274435/35.  
 DR P-PSDB; AAR42280.  
 XX  
 PT New prostaglandin E receptor protein and DNA encoding it - used to study  
 PT prostaglandin binding, agonists and antagonists etc.  
 XX  
 PS Claim 15; Fig 1; 22pp; English.  
 XX  
 CC PGE2 is known to be involved in digestive tract constriction/ relaxation,  
 CC gastric juice secretion and neurotransmitter release. The protein encoded  
 CC by the gene is capable of receiving PGE and of activating GTP-binding  
 CC protein. (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 2107 BP; 482 A; 549 C; 563 G; 513 T; 0 U; 0 Other;  
 Query Match 78.9%; Score 15; DB 2; Length 2107;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGTAGCGGCCCGCGC 15  
 Db 644 TGTAGCGGCCCGCGC 630  
 RESULT 13  
 ABV75863/c  
 ID ABV75863 standard; DNA; 14800 BP.  
 XX  
 AC ABV75863;  
 XX  
 DT 05-FEB-2003 (first entry)  
 XX  
 DE Human potassium channel transporter gene.  
 XX  
 KW Transporter; potassium channel; human; gene therapy; chromosome 17; gene;  
 KW ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 2374..12713  
 FT /tag= a  
 FT /product= "Transporter protein"  
 FT /note= "contains introns"  
 FT 2374..2449  
 FT /tag= b  
 FT /number= 1  
 FT 2450..4293  
 FT /tag= c  
 FT /number= 1  
 FT 4294..4527  
 FT /tag= d  
 FT /number= 2  
 FT 4528..5092  
 FT /tag= e  
 FT /number= 2  
 FT 5093..5220  
 FT /tag= f  
 FT exon  
 FT intron

FT /number= 3  
FT 5221. .7022  
FT /\*tag= 9  
FT /number= 3  
FT 7023. .7266  
FT /\*tag= h  
FT /number= 4  
FT 7267. .7583  
FT /\*tag= i  
FT /number= 4  
FT 7584. .7741  
FT /\*tag= j  
FT /number= 5  
FT 7742. .11326  
FT /\*tag= k  
FT /number= 5  
FT replace(10738,GTG)  
FT /\*tag= 0  
FT /standard\_name= "Single nucleotide polymorphism"  
FT 11327. .11534  
FT /\*tag= 1  
FT /number= 6  
FT 11535. .12468  
FT /\*tag= m  
FT /number= 6  
FT 12469. .12710  
FT /\*tag= n  
FT /number= 7  
FT exon  
FT 30-MAR-2001; 2001US-00820923.  
FT PR  
FT WO200281657-A2.  
FT PN  
FT 17-OCT-2002.  
FT PD  
FT 01-APR-2002; 2002WO-US009743.  
FT PF  
FT 30-MAR-2001; 2001US-00820923.  
FT PR  
FT (PEKE ) PE CORP NY.  
FT (BEAS/) BEASLEY E M.  
FT (WEIM/) WEI M.  
FT (KERC/) KETCHUM K A.  
FT (DFRA/) DI FRANCESCO V.  
FT PA  
FT Beasley EM, Wei M, Ketchum KA, Di Francesco V;  
FT PI  
FT WPI; 2003-058533/05.  
FT DR  
FT P-PSDB; ABP55112.  
FT DR  
FT Novel isolated human transporter peptide useful for treating disorder  
FT PT characterized by absence of, in appropriate or unwanted expression of the  
FT PT transporter protein, and as immunogens to raise antibodies.  
FT XX  
FT Claim 4; Fig 3A-D; 75pp; English.  
FT PS  
XX The present sequence is that of genomic DNA encoding a novel human  
CC transporter protein that is related to the voltage gated potassium  
CC channel transporter family, and which is expressed in the lung and in the  
CC foetal brain. The gene is located on chromosome 17. Nucleic acids  
CC encoding the transporter are useful as probes, primers, in biological  
CC assays, for constructing recombinant vectors, host cells and transgenic  
CC animals, for expressing antigenic portions of the transporter protein,  
CC for designing ribozymes, for monitoring the effectiveness of modulating  
CC compounds on the expression or activity of the human transporter gene in  
CC clinical trials or in a treatment regimen, in diagnostic assays for  
CC qualitative changes in human transporter nucleic acid that lead to  
CC pathology, for testing an individual for a genotype that while not  
CC necessarily causing a disease, nevertheless affects the treatment  
CC modality, as antisense constructs to control human transporter gene  
CC expression in cells, tissues and organisms, and for gene therapy in  
CC patients containing cells that are aberrant in human transporter gene  
CC expression. A gene chip comprising a transporter nucleic acid is claimed,  
CC and is useful for conducting cell-based assays involving the human  
CC transporter protein

XX SQ Sequence 14800 BP; 3445 A; 4202 C; 3942 G; 3211 T; 0 U; 0 Other;  
SQ Query Match 78.9%; Score 15; DB 10; Length 14800;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 4 AGCGGCCGACGCGTG 18  
DB 2050 AGCGGCCGACGCGTG 2036  
RESULT 14  
ACF04818  
ID ACF04818 standard; DNA; 51855 BP.  
XX  
XX ACF04818;  
XX AC  
XX 12-FEB-2004 (first entry)  
XX DT  
XX DE Melithiazol biosynthetic gene cluster.  
XX DE  
XX Melithiazol synthetase; polyketide synthase; fungicide;  
XX KM nonribosomal peptide synthetase; gene; ds.  
XX KM  
XX OS  
XX Unidentified.  
XX OS  
XX WO2003080828-A2.  
XX PN  
XX 02-OCT-2003.  
XX PD  
XX 21-MAR-2003; 2003WO-EP002387.  
XX PF  
XX 21-MAR-2002; 2002EP-00006389.  
XX PR  
XX (GBFB ) GBF GBS BIOTECH FORSCHUNG GMBH.  
XX PA  
XX Mueller R, Weinig S, Bloecker H, Hoeffle G, Saase F;  
XX PI  
XX WPI; 2003-767886/72.  
XX DR  
XX New nucleic acid molecule termed mel genes encoding polypeptide synthase  
FT PT and non-ribosomal peptide synthetase that are involved in synthesis of  
FT PT melithiazol displaying enhanced fungicidal activity.  
XX PT  
XX Disclosure; Fig 2; 0pp; English.  
XX PS  
CC The present invention provides the protein and coding sequences of a  
CC number of proteins involved in the biosynthesis of melithiazol. These are  
CC designated Mel B - Mel H and ORF 1 - ORF 7. The sequences can be used in  
CC the production of melithiazol, which exhibits fungicidal activity. The  
CC present sequence is a coding sequence shown in the exemplification of the  
CC invention  
XX CC  
XX SQ Sequence 51855 BP; 7256 A; 17204 C; 19038 G; 8357 T; 0 U; 0 Other;  
SQ Query Match 78.9%; Score 15; DB 10; Length 51855;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TGTAGCGGCCGACGC 15  
DB 3329 TGTAGCGGCCGACGC 3343  
RESULT 15  
ADL13471/c  
ID ADL13471 standard; DNA; 214520 BP.  
XX  
XX ADL13471;  
XX AC  
XX 06-MAY-2004 (first entry)  
XX DT  
XX

```

DE Osteoarthritis-associated polymorphic nucleotide #3.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
OS
XX WO2003054166-A2.
PN
XX
XX 03-JUL-2003.
PD
XX
XX 19-DEC-2002; 2002WO-US041225.
PF
XX
XX 20-DEC-2001; 2001US-0342603P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Jones KA, Schafer A;
PI
XX
XX WPI; 2003-559141/52.
DR
XX
XX
XX Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX
XX Disclosure; SEQ ID NO 3; 297pp; English.
PS
XX
XX The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC the sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX
XX Sequence 214520 BP; 57899 A; 52039 C; 51001 G; 53577 T; 0 U; 4 Other;
SQ
Query Match 78.9%; Score 15; DB 10; Length 214520;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTAGCGGCCACCGC 15
DB 120357 TGTAGCGGCCACCGC 120343

```

Search completed: February 10, 2005, 13:44:04  
 Job time : 22.6391 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 / Search time 153.034 Seconds  
(without alignment)  
5324.730 Million cell updates/sec

Title: US-10-790-430-7

Sequence: 1 aatcgatccaaatcgcgacg.....aaaggtgcccgaagtcaa 498

Scoring table: OLIGO NUC  
Gapop 60.0, Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCITUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 498   | 100.0       | 498    | 4     | US-09-872-051-7     |
| 2          | 118   | 23.7        | 1183   | 4     | US-09-872-051-8     |
| 3          | 101   | 20.3        | 2378   | 3     | US-09-441-340-27    |
| 4          | 70    | 14.1        | 1404   | 1     | US-08-144-6028-7    |
| 5          | 70    | 14.1        | 2159   | 1     | US-08-144-6028-5    |
| 6          | 70    | 14.1        | 5643   | 1     | US-08-144-6028-4    |
| 7          | 68    | 13.7        | 1392   | 1     | US-08-144-6028-6    |
| 8          | 67    | 13.5        | 4032   | 3     | US-09-068-101-5     |
| 9          | 67    | 13.5        | 4032   | 4     | US-09-970-921-5     |
| 10         | 63    | 12.7        | 399    | 4     | US-09-682-597A-5    |
| 11         | 30    | 6.0         | 30     | 4     | US-09-682-597A-2    |
| 12         | 29    | 5.8         | 29     | 4     | US-09-872-051-4     |
| 13         | 22    | 4.4         | 22     | 4     | US-09-872-051-13    |
| 14         | 22    | 4.4         | 22     | 4     | US-09-872-051-14    |
| 15         | 20    | 4.0         | 2107   | 3     | US-09-441-340-29    |
| 16         | 20    | 4.0         | 2436   | 3     | US-09-441-340-31    |
| 17         | 20    | 4.0         | 3039   | 4     | US-09-377-466B-19   |
| 18         | 20    | 4.0         | 3039   | 4     | US-09-377-466B-21   |
| 19         | 20    | 4.0         | 3044   | 4     | US-09-377-466B-38   |
| 20         | 20    | 4.0         | 3450   | 4     | US-09-377-466B-17   |
| 21         | 20    | 4.0         | 3455   | 4     | US-09-377-466B-36   |
| 22         | 20    | 4.0         | 3469   | 4     | US-09-377-466B-23   |
| 23         | 20    | 4.0         | 3754   | 4     | US-09-377-466B-13   |
| 24         | 20    | 4.0         | 4149   | 4     | US-09-377-466B-13   |
| 25         | 20    | 4.0         | 8349   | 4     | US-09-166-002-16    |
| 26         | 20    | 4.0         | 254778 | 4     | US-09-949-016-12417 |
| 27         | 19    | 3.8         | 19     | 4     | US-09-872-051-9     |

|   |    |    |     |        |   |                      |                   |
|---|----|----|-----|--------|---|----------------------|-------------------|
| c | 28 | 19 | 3.8 | 601    | 4 | US-09-949-016-192515 | Sequence 192515,  |
| c | 29 | 19 | 3.8 | 1239   | 3 | US-09-180-342-1      | Sequence 1, Appl1 |
| c | 30 | 19 | 3.8 | 187133 | 4 | US-09-949-016-17231  | Sequence 17231, A |
| c | 31 | 18 | 3.6 | 40     | 1 | US-08-368-803-8      | Sequence 8, Appl1 |
| c | 32 | 18 | 3.6 | 40     | 1 | US-08-368-803-9      | Sequence 9, Appl1 |
| c | 33 | 18 | 3.6 | 601    | 4 | US-09-949-016-163649 | Sequence 163649,  |
| c | 34 | 18 | 3.6 | 601    | 4 | US-09-949-016-163650 | Sequence 163650,  |
| c | 35 | 18 | 3.6 | 601    | 4 | US-09-949-016-165157 | Sequence 165157,  |
| c | 36 | 18 | 3.6 | 601    | 4 | US-09-949-016-180872 | Sequence 180872,  |
| c | 37 | 18 | 3.6 | 601    | 4 | US-09-949-016-180873 | Sequence 180873,  |
| c | 38 | 18 | 3.6 | 601    | 4 | US-09-949-016-205828 | Sequence 205828,  |
| c | 39 | 18 | 3.6 | 1038   | 4 | US-09-248-796A-4729  | Sequence 4729, Ap |
| c | 40 | 18 | 3.6 | 1419   | 3 | US-08-174-467-17     | Sequence 17, Appl |
| c | 41 | 18 | 3.6 | 1419   | 3 | US-08-452-071-17     | Sequence 17, Appl |
| c | 42 | 18 | 3.6 | 2463   | 4 | US-09-573-080A-18    | Sequence 18, Appl |
| c | 43 | 18 | 3.6 | 5000   | 4 | US-09-949-016-452    | Sequence 452, App |
| c | 44 | 18 | 3.6 | 5001   | 4 | US-09-949-016-1315   | Sequence 1315, Ap |
| c | 45 | 18 | 3.6 | 9953   | 4 | US-09-949-016-15190  | Sequence 15190, A |

#### ALIGNMENTS

```
RESULT 1
US-09-872-051-7
Sequence 7, Application US/09872051
Patent No. 6825400
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Behr, Carl
APPLICANT: Hironaka, Catherine
APPLICANT: Heck, Gregory
APPLICANT: You, Jinsong
TITLE OF INVENTION: Corn Event PV-ZMG732 (mk603) and Composition and Methods for Det
FILE REFERENCE: 38-21(52258)B
CURRENT APPLICATION NUMBER: US/09/872,051
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/213,567
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/241,215
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/240,014
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 498
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: source
LOCATION: (1)..(498)
OTHER INFORMATION: 1-304 Zea maize genomic DNA
OTHER INFORMATION: 305-349 construct vector DNA
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA
US-09-872-051-7
Query Match 100.0%; Score 498; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.6e-244;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | AATGATCCAAATCGGACTGGAATGTGGAGAAAGAGAGAAACAGAGAGCCCTCAGGTT | 60  |
| DB | 1   | AATGATCCAAATCGGACTGGAATGTGGAGAAAGAGAGAAACAGAGAGCCCTCAGGTT | 60  |
| QY | 61  | TCGAGGTGAAGTATCAGAGATTACCGCCATGCTTTATGAGACAAGAGGGGAG      | 120 |
| DB | 61  | TCGAGGTGAAGTATCAGAGATTACCGCCATGCTTTATGAGACAAGAGGGGAG      | 120 |
| QY | 121 | GAGGTAAACAGATCAGATCAGCGCTGGAAGTTTCTCAAGAGATCGGAAGCTGTTCC  | 180 |
| DB | 121 | GAGGTAAACAGATCAGATCAGCGCTGGAAGTTTCTCAAGAGATCGGAAGCTGTTCC  | 180 |

|    |     |   |       |
|----|-----|---|-------|
| OY | 181 | AGCGCCGCGTGGCATTTCGGCGAATCCTCTCTCTCGGGATAGAGCCGATCTTTCTCTG    | 24 48 |
| Db | 181 | AGCGCCGCGTGGCATTTCGGCGAATCCTCTCTCTCGGGATAGAGCCGATCTTTCTCTG    | 24 48 |
| OY | 241 | GCAATTCCAAACCTTAGAGACGTGCGTCCCTGGTGGGCTGCTCGGCCAGCAAGCCTTGTAG | 3 00  |
| Db | 241 | GCAATTCCAAACCTTAGAGACGTGCGTCCCTGGTGGGCTGCTCGGCCAGCAAGCCTTGTAG | 3 00  |
| OY | 301 | CGGCCACGCGTGGTACCAAGCTTGATATCCCTAGGGCGGCGCGTTAACAGCTTACTC     | 3 60  |
| Db | 301 | CGGCCACGCGTGGTACCAAGCTTGATATCCCTAGGGCGGCGCGTTAACAGCTTACTC     | 3 60  |
| OY | 361 | GAGGTCAATTCATATGCTTGAGAAAGAGAGTGGGATAGTCCAAATATATAACAAAGGTAGA | 4 20  |
| Db | 361 | GAGGTCAATTCATATGCTTGAGAAAGAGAGTGGGATAGTCCAAATATATAACAAAGGTAGA | 4 20  |
| OY | 421 | TTACCGGTCAAAAGTGAACAATCAGTTAAAGGTGTATTAAGTAATAATTCGGTATAAA    | 4 80  |
| Db | 421 | TTACCGGTCAAAAGTGAACAATCAGTTAAAGGTGTATTAAGTAATAATTCGGTATAAA    | 4 80  |
| OY | 481 | AAGGTGGCCCAAAAGTGA 498  |       |
| Db | 481 | AAGGTGGCCCAAAAGTGA 498  |       |

## RESULT 2

```

US-09-872-051-8/c
; Sequence 8, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZWGT32(nk603) and Composition and Methods for Detecting and Compositing
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1) .. (1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 382-686 Zea maize plastid genes, tpell and tpoa
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
; US-09-872-051-8

```

|                       |              |   |          |             |
|-----------------------|--------------|---|----------|-------------|
| Query Match           | 23.7%        | Score 118   | DB 4     | Length 1183 |
| Best Local Similarity | 100.0%       | Pred. No.   | 2.7e-50  |             |
| Matches 118           | Conservative | Mismatches 0  | Indels 0 | Gaps 0      |
| QY                    | 308          | CGCGTGTGAACCAAGCTTGATATCCCTAGGGCGCGCCCGTTAACAGACTTACTCGAGGTCA | 367      |             |
| Db                    | 381          | CGCGGTGAACCAAGCTTGATATCCCTAGGGCGCGCCCGTTAACAGACTTACTCGAGGTCA  | 322      |             |
| QY                    | 368          | TTCATATCTTGAGAAGAGTGGGATATGTCACAAATATAAACAAAGGTATAGATTACC     | 425      |             |
| Db                    | 321          | TTCATATCTTGAGAAGAGTGGGATATGTCACAAATATAAACAAAGGTATAGATTACC     | 264      |             |

```

RESULT 3
US-09-441-340-27
; Sequence 27, Application US/09441340
; Patent No. 6448476
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21 (15303)
; CURRENT APPLICATION NUMBER: US/09/441,340
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 60/108,763
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27

```

## FEAT

```

? OTHER INFORMATION: Description of Artificial Sequence: expression
? OTHER INFORMATION: cassette comprising a plant promoter linked to an
? OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
? OTHER INFORMATION: transferase, and a termination sequence
? FEATURE:
? NAME/KEY: promoter
? LOCATION: (28)..(965)
? FEATURE:
? NAME/KEY: intron
? LOCATION: (966)..(1423)
? FEATURE:
? NAME/KEY: translt peptide
? LOCATION: (1440)..(1667)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1668)..(2099)
? FEATURE:
? NAME/KEY: terminator
? LOCATION: (2114)..(2369)
US-09-441-340-27

```

|                           |        |                    |       |              |
|---------------------------|--------|--------------------|-------|--------------|
| Query Match               | 20.3%  | Score 101;         | DB 3; | Length 2378; |
| Best Local Similarity     | 100.0% | Pred. No. 1.3e-41; |       |              |
| Matches 101; Conservative | 0;     | Mismatches         | 0;    | Gaps 0       |

QY 325 GATATCCCTTGGGGGGGGCGGGTTAACAGAGTTACTCGAGTCATTCAATGCTTGAAG 38  
Db 1 GATATCCCTTGGGGGGGGCGGGTTAACAGAGTTACTCGAGTCATTCAATGCTTGAAG 60  
QY 385 AGAGTCGGGATAGTCCAAATAAACAAGTAGATTACC 425  
Db 61 AGAGTCGGGATAGTCCAAATAAACAAGTAGATTACC 101

## RESULT 4

```

US-08-144-602B-7
: Sequence 7, Application US/08144602B
: Patent No. 5641876
:
: GENERAL INFORMATION:
:
: APPLICANT: MCELROY, David
:
: APPLICANT: WU, Ray
:
: TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
:
: NUMBER OF SEQUENCES: 27
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
:
: STREET: CLINTON SQUARE, P.O. BOX 1051
:
: CITY: ROCHESTER
:
: STATE: NEW YORK
:
: COUNTRY: USA
:
: ZIP: 14603
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/144,602B  
FILING DATE: 27-OCT-1993  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/10140  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-144-602B-7

Query Match 14.1%; Score 70; DB 1; Length 1404;  
Best Local Similarity 100.0%; Pred. No. 9.2e-26;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAAACAAGG 415  
DB 9 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAAACAAGG 68

QY 416 TAAATTACC 425  
DB 69 TAAATTACC 78

RESULT 5  
US-08-144-602B-5  
Sequence 5, Application US/08144602B  
Patent No. 5641876  
GENERAL INFORMATION:  
APPLICANT: McElroy, David  
TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: CLINTON SQUARE, P.O. BOX 1051  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/144,602B  
FILING DATE: 27-OCT-1993  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/10140  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2199 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-144-602B-5

Query Match 14.1%; Score 70; DB 1; Length 2199;  
Best Local Similarity 100.0%; Pred. No. 9.1e-26;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAAACAAGG 415  
DB 806 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAAACAAGG 865

QY 416 TAAATTACC 425  
DB 866 TAAATTACC 875

RESULT 6  
US-08-144-602B-4  
Sequence 4, Application US/08144602B  
Patent No. 5641876  
GENERAL INFORMATION:  
APPLICANT: McElroy, David  
TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: CLINTON SQUARE, P.O. BOX 1051  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/144,602B  
FILING DATE: 27-OCT-1993  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/10140  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-144-602B-4

Query Match 14.1%; Score 70; DB 1; Length 5643;  
Best Local Similarity 100.0%; Pred. No. 9e-26;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAAACAAGG 415  
DB 809 TACTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATAGTCCAAATATAAACAAGG 868

QY 416 TAAATTACC 425  
DB 869 TAAATTACC 878

RESULT 7  
US-08-144-602B-6

```

Sequence 6, Application US/08144602B
Patent No. 5641876
GENERAL INFORMATION:
APPLICANT: McElroy, David
TITLE OF INVENTION: RICE ACTIN GENE AND PROMOTER
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: CLINTON SQUARE, P.O. BOX 1051
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/144,602B
FILING DATE: 27-OCT-1993
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/10140
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-144-602B-6

```

```

Query Match      13.7%; Score 68; DB 1; Length 1392;
Best Local Similarity 100.0%; Pred. No. 9.7e-25;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      358 CTCGAGGTCATTGCTATGCTTGAGAGAGAGTCGGGATAGTCCAAATATTAACCAAGGTA 417
Db      1 CTCGAGGTCATTGCTATGCTTGAGAGAGAGTCGGGATAGTCCAAATATTAACCAAGGTA 60

Qy      418 AGATTACC 425
Db      61 AGATTACC 68

RESULT 8
US-09-068-101-5
Sequence 5, Application US/09068101
GENERAL INFORMATION:
APPLICANT: PLANT GENETIC SYSTEMS N.V.
TITLE OF INVENTION: Improved Barstar Gene
FILE REFERENCE: 2121-139P
CURRENT APPLICATION NUMBER: US/09/068,101
EARLIER FILING DATE: 1998-08-26
EARLIER APPLICATION NUMBER: EP 96202446.9
EARLIER FILING DATE: 1996-09-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 4032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: other nucleic
OTHER INFORMATION: acid, "plasmid pmv71"

```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1999)..(3400)
OTHER INFORMATION: label = PRAc1, "promoter region of rice actin gene"
OTHER INFORMATION: - contains an intron in the leader"
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3401)..(3676)
OTHER INFORMATION: label = barstar, "barstar DNA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3677)..(4003)
OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium"
OTHER INFORMATION: T-DNA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3395)..(3404)
OTHER INFORMATION: label = NcoI, "NcoI recognition site"
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4015)..(4021)
OTHER INFORMATION: label = KpnI, "KpnI recognition site"
US-09-068-101-5

```

```

Query Match      13.5%; Score 67; DB 3; Length 4032;
Best Local Similarity 100.0%; Pred. No. 3.1e-24;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      359 TCGAGGTCATTGCTATGCTTGAGAGAGAGTCGGGATAGTCCAAATATTAACCAAGGTA 418
Db      1999 TCGAGGTCATTGCTATGCTTGAGAGAGAGTCGGGATAGTCCAAATATTAACCAAGGTA 2058

Qy      419 GATTACC 425
Db      2059 GATTACC 2065

```

```

RESULT 9
US-09-970-921-5
Sequence 5, Application US/09970921
Patent No. 6759575
GENERAL INFORMATION:
APPLICANT: Frank Michiels et al.
TITLE OF INVENTION: Improved Barstar Gene
FILE REFERENCE: 2428-0108P
CURRENT APPLICATION NUMBER: US/09/970,921
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 4032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: other nucleic
OTHER INFORMATION: acid, "plasmid pmv71"
NAME/KEY: misc_feature
LOCATION: (1995)..(3400)
OTHER INFORMATION: label = PRAc1, "promoter region of rice actin gene"
OTHER INFORMATION: - contains an intron in the leader"
NAME/KEY: misc_feature
LOCATION: (3401)..(3676)
OTHER INFORMATION: label = barstar, "barstar DNA"
NAME/KEY: misc_feature
LOCATION: (3677)..(4003)
OTHER INFORMATION: label = 3'nos, "region containing 3' untranslated
OTHER INFORMATION: end of nopaline synthase gene of Agrobacterium"
OTHER INFORMATION: T-DNA"
NAME/KEY: misc_feature
LOCATION: (3395)..(3404)
OTHER INFORMATION: label = NcoI, "NcoI recognition site"
NAME/KEY: misc_feature

```

```
; LOCATION: (4016)..(4021)
; OTHER INFORMATION: label = Kpnl, "Kpnl recognition site"
US-09-970-921-5

Query Match
Best Local Similarity 13.5%; Score 67; DB 4; Length 4032;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 359 TCGAGGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATTAACCAAGGTAA 418
DB 1999 TCGAGGTCATTGATGCTTGAGAGAGAGTCGGATAGTCCAAATTAACCAAGGTAA 2058

OY 419 GATTACC 425
DB 2059 GATTACC 2065

RESULT 10
US-09-682-597A-5
; Sequence 5, Application US/09682597A
; Patent No. 6689880
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Chen, Guilan
; APPLICANT: Hironaka, Catherine
; APPLICANT: Zhou, Hua-Ping
; TITLE OF INVENTION: Glyphosate Tolerant Wheat Plant 33391 and Compositions and Methods
; FILE REFERENCE: 38-21(52232)A
; CURRENT APPLICATION NUMBER: US/09/682,597A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: m1ec feature
; LOCATION: (1)..(399)
; OTHER INFORMATION: chimeric sequence of wheat genome and transgene insert
US-09-682-597A-5

Query Match
Best Local Similarity 12.7%; Score 63; DB 4; Length 399;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 337 GCGGCGCGTTACACAGCTTACTGAGTCATTGATGCTTGAGAGAGAGTCGGATA 396
DB 337 GCGGCGCGTTACACAGCTTACTGAGTCATTGATGCTTGAGAGAGAGTCGGATA 396

OY 397 GTC 399
DB 397 GTC 399

RESULT 11
US-09-682-597A-2/c
; Sequence 2, Application US/09682597A
; Patent No. 6689880
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Chen, Guilan
; APPLICANT: Hironaka, Catherine
; APPLICANT: Zhou, Hua-Ping
; TITLE OF INVENTION: Glyphosate Tolerant Wheat Plant 33391 and Compositions and Methods
; FILE REFERENCE: 38-21(52232)A
; CURRENT APPLICATION NUMBER: US/09/682,597A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
```

```
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-682-597A-2

Query Match
Best Local Similarity 6.0%; Score 30; DB 4; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 370 CATATGCTTGAGAGAGAGTCGGATAGTC 399
DB 30 CATATGCTTGAGAGAGAGTCGGATAGTC 1

RESULT 12
US-09-872-051-4/c
; Sequence 4, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGR32 (nke03) and Composition and Methods for Detection
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(29)
; OTHER INFORMATION: fully synthesized
US-09-872-051-4

Query Match
Best Local Similarity 5.8%; Score 29; DB 4; Length 29;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 GAGTCGGATAGTCACAAATTAACCAAG 414
DB 29 GAGTCGGATAGTCACAAATTAACCAAG 1

RESULT 13
US-09-872-051-13
; Sequence 13, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGR32 (nke03) and Composition and Methods for Detection
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
```

```

; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(22)
; OTHER INFORMATION: fully synthesized
US-09-872-051-13

Query Match          4.4%; Score 22; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AATGATCCAAATCGGACTG 22
DB      1 AATGATCCAAATCGGACTG 22

RESULT 14
US-09-872-051-14/c
; Sequence 14, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMG732 (n603) and Composition and Methods for Detect
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(22)
; OTHER INFORMATION: fully synthesized
US-09-872-051-14

Query Match          4.4%; Score 22; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      477 ATAAAGGTGGCCCAAGTGAA 498
DB      22 ATAAAGGTGGCCCAAGTGAA 1

RESULT 15
US-09-441-340-29
; Sequence 29, Application US/09441340
; Patent No. 6448476
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
```

```

; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/09/441,340
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 60/108,763
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 2107
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Description of Artificial Sequence: expression
; OTHER INFORMATION: cassette comprising plant operable promoter linked
; OTHER INFORMATION: to a leader, intron, a sequence encoding an AMPA
; OTHER INFORMATION: acetyltransferase, and termination sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (26)..(590)
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (615)..(685)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (699)..(1148)
; FEATURE:
; NAME/KEY: transic peptide
; LOCATION: (1149)..(1426)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1427)..(1858)
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (1869)..(2102)
US-09-441-340-29

Query Match          4.0%; Score 20; DB 3; Length 2107;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      337 GCGGCGCGTTAACACACTT 356
DB      1 GCGGCGCGTTAACACACTT 20

Search completed: February 10, 2005, 08:49:45
Job time : 154.034 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 1222.79 Seconds  
(without alignments)  
5727.097 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183  
Sequence: 1 gacgttatcatgatgagatggg.....aaggaagcgcgagatgacg 1183

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: Geneseq\_16Dec04.\*  
2: geneseq1980s.\*  
3: geneseq1990s.\*  
4: geneseq2000s.\*  
5: geneseq2001as.\*  
6: geneseq2002as.\*  
7: geneseq2002bs.\*  
8: geneseq2003as.\*  
9: geneseq2003bs.\*  
10: geneseq2003cs.\*  
11: geneseq2003ds.\*  
12: geneseq2004as.\*  
13: geneseq2004bs.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 1183  | 100.0       | 1183   | ABK15245 | Abk15245 Corn nk60  |
| 2          | 202   | 17.1        | 2378   | AA001016 | AA001016 Expressio  |
| 3          | 171   | 14.5        | 1501   | AA087194 | AA087194 Rice acti  |
| 4          | 171   | 14.5        | 1501   | AA089325 | AA089325 Rice acti  |
| 5          | 171   | 14.5        | 2181   | AD084543 | AD084543 Rice acti  |
| 6          | 171   | 14.5        | 9359   | ABK11039 | ABK11039 PVDH636 v  |
| 7          | 171   | 14.5        | 9359   | ABK10687 | ABK10687 Transform  |
| 8          | 169   | 14.3        | 5118   | ADP79330 | ADP79330 Plasmid P  |
| 9          | 169   | 14.3        | 7545   | ACF58343 | ACF58343 Nucleotid  |
| 10         | 169   | 14.3        | 11643  | ACF58344 | ACF58344 Nucleotid  |
| 11         | 168   | 14.2        | 4032   | AAV23237 | AAV23237 Plasmid P  |
| 12         | 163   | 13.8        | 6865   | AA080287 | AA080287 Nucleotid  |
| 13         | 163   | 13.8        | 10003  | AA080295 | AA080295 Nucleotid  |
| 14         | 159   | 13.4        | 4774   | ADP83768 | ADP83768 Benthicre  |
| 15         | 147   | 12.4        | 349    | AA051491 | AA051491 A. tumefa  |
| 16         | 147   | 12.4        | 896    | ADP83767 | ADP83767 Benthicre  |
| 17         | 147   | 12.4        | 1360   | AA051480 | AA051480 Consensus  |
| 18         | 147   | 12.4        | 3754   | AA051642 | AA051642 Bacillus   |
| 19         | 147   | 12.4        | 3754   | AA061790 | AA061790 Bt. cry3Bb |
| 20         | 147   | 12.4        | 3754   | ABX95186 | ABX95186 B. thurin  |

#### ALIGNMENTS

|    |     |      |       |    |          |                     |
|----|-----|------|-------|----|----------|---------------------|
| 21 | 147 | 12.4 | 4149  | 3  | AA051641 | AA051641 Bacillus   |
| 22 | 147 | 12.4 | 4149  | 10 | AA061789 | AA061789 Bt. cry3Bb |
| 23 | 147 | 12.4 | 4149  | 10 | ABX95185 | ABX95185 B. thurin  |
| 24 | 147 | 12.4 | 8349  | 3  | AA015565 | AA015565 pMON30464  |
| 25 | 147 | 12.4 | 10249 | 3  | AA015563 | AA015563 pMON33828  |
| 26 | 147 | 12.4 | 10312 | 3  | AA015564 | AA015564 pMON33829  |
| 27 | 147 | 12.4 | 10339 | 3  | AA015566 | AA015566 pMON33827  |
| 28 | 144 | 12.2 | 6600  | 12 | ADK98482 | ADK98482 B. thurin  |
| 29 | 144 | 12.2 | 7000  | 12 | ADK98485 | ADK98485 B. thurin  |
| 30 | 140 | 11.8 | 2378  | 3  | AA01016  | AA01016 Expressio   |
| 31 | 140 | 11.8 | 2436  | 3  | AA01018  | AA01018 Expressio   |
| 32 | 137 | 11.6 | 1630  | 3  | AA01014  | AA01014 Expressio   |
| 33 | 137 | 11.6 | 6313  | 6  | ABA05258 | ABA05258 Human ICA  |
| 34 | 137 | 11.6 | 6313  | 10 | AD097340 | AD097340 Immunoadh  |
| 35 | 137 | 11.6 | 8074  | 6  | ABA05262 | ABA05262 Plasmid P  |
| 36 | 137 | 11.6 | 8074  | 10 | AD097345 | AD097345 Immunoadh  |
| 37 | 137 | 11.6 | 10846 | 6  | AB054336 | AB054336 E. coli f  |
| 38 | 137 | 11.6 | 10847 | 2  | AA008923 | AA008923 Vector co  |
| 39 | 137 | 11.6 | 10900 | 2  | AA008924 | AA008924 Vector co  |
| 40 | 137 | 11.6 | 10900 | 6  | AB054337 | AB054337 E. coli f  |
| 41 | 137 | 11.6 | 11606 | 12 | AD013598 | AD013598 Plasmid P  |
| 42 | 136 | 11.5 | 1691  | 4  | AA019910 | AA019910 pG1 vecto  |
| 43 | 136 | 11.5 | 1851  | 4  | AA019912 | AA019912 pG1 vecto  |
| 44 | 136 | 11.5 | 2315  | 4  | AA019911 | AA019911 pG16 vect  |
| 45 | 136 | 11.5 | 2708  | 4  | AA019908 | AA019908 pTBN6 vec  |

RESULT 1  
ABK15245  
ID ABK15245 standard; DNA, 1183 BP.  
XX  
AC ABK15245;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Corn nk603 3' transgene/genomic sequence PCR fragment.  
XX  
KW ds; rice actin 1 promoter; RA1; RA1 intron;  
KW chloroplast transit peptide gene; glycosylase resistance; corn;  
KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
KW transcripional terminator; cauliflower mosaic virus 35S promoter;  
KW PV-2WGT32; transgenic; nk603.  
XX  
XX Zea mays.  
OS Escherichia coli.  
OS Agrobacterium tumefaciens.  
OS Synthetic.  
OS Chimeric.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1..164  
FT FT /\*tag= a  
FT FT /note= "T-AGRTU.nos vector sequence"  
FT FT misc\_feature 165..381  
FT FT /\*tag= b  
FT FT /note= "Vector sequence"  
FT FT misc\_feature 382..686  
FT FT /\*tag= c  
FT FT /note= "Corn plasmid genes rps11/rpoA partial sequence"  
FT FT misc\_feature 687..1183  
FT FT /\*tag= d  
FT FT /note= "Corn genomic sequence"  
PN BP167531-AI.  
XX  
XX 02-JAN-2002.  
XX  
XX 15-JUN-2001; 2001EP-002023314.  
XX PF AA051642 Bacillus  
XX  
XX AA061790 Bt. cry3Bb  
XX  
XX 22-JUN-2000; 2000US-0213567P.

PR 13-OCT-2000; 2000US-0240014P.  
 PR 13-OCT-2000; 2000US-0241215P.  
 XX  
 PA (MONS ) MONSANTO TECHNOLOGY LLC.

PI Behr CF, Hironaka C, Heck GR, You J;  
 DR WPI; 2002-165871/22.

XX Novel DNA construct useful for producing a corn plant that tolerates  
 PT application of glyphosate herbicide, comprises two transgene expression  
 PT cassettes.

XX Claim 3; Page 16-17; 25pp; English.

XX The invention relates to a DNA construct (I) comprising a first (F1) and  
 CC a second (F2) expression cassette. F1 of the DNA construct, in operable  
 CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
 CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
 CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
 CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
 CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
 CC P1-ZMGT32(hk603) of 498 or 1183bp, or primers and probes derived from  
 CC them (used to detect transgene junction points). The construct is useful  
 CC for producing a corn plant that tolerates application of glyphosate  
 CC herbicide by transforming a corn cell with, selecting the corn cell for  
 CC tolerance to application of glyphosate, and growing the corn cell into a  
 CC fertile corn plant. The present sequence is the 1183bp amplicon  
 CC consisting of the Agrobacterium transcriptional terminator sequence,  
 CC vector sequences and corn sequences

XX Sequence 1183 BP; 295 A; 289 G; 280 G; 319 T; 0 U; 0 Other;

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTTATTTAGATGCGTTTATGATTTAGAGTCCCGCAATTATACATTTAATACGC 60  
 DB 1 GAGCTTTATTTAGATGCGTTTATGATTTAGAGTCCCGCAATTATACATTTAATACGC 60  
 QY 61 GATGAAAAAATAATATGCGCGCAACTAGATTAATTATTCGCGCGGTCTATCTAT 120  
 DB 61 GATGAAAAAATAATATGCGCGCAACTAGATTAATTATTCGCGCGGTCTATCTAT 120  
 QY 121 GTTACTGATGCGGGAATATCCCGGGAATTCGATACCAAGCTTTATATAGTAGAATA 180  
 DB 121 GTTACTGATGCGGGAATATCCCGGGAATTCGATACCAAGCTTTATATAGTAGAATA 180  
 QY 181 GAGTAAATTTACATTTGGGCGCACTTTTATTAACGATATTTTACTTTATACACCTTTTA 240  
 DB 181 GAGTAAATTTACATTTGGGCGCACTTTTATTAACGATATTTTACTTTATACACCTTTTA 240  
 QY 241 ACTGATGTTTCACTTTGACCAAGTATCTTACTTTGTTTATTTTGACATCCGCA 300  
 DB 241 ACTGATGTTTCACTTTGACCAAGTATCTTACTTTGTTTATTTTGACATCCGCA 300  
 QY 301 CTCTCTCTCAAGATATGATGACCTCGAGTAAGCTTTTAAAGCGCGCCCTTAGGGA 360  
 DB 301 CTCTCTCTCAAGATATGATGACCTCGAGTAAGCTTTTAAAGCGCGCCCTTAGGGA 360  
 QY 361 TATCAAGCTTGTGACACGACGACACACTTCACTTACTTATGTTGATGATCTGTTATC 420  
 DB 361 TATCAAGCTTGTGACACGACGACACACTTCACTTACTTATGTTGATGATCTGTTATC 420  
 QY 421 TCTTCTGAAACATPAACAGACTAGTATTTGATCATTTGATGATGATCTTCTTAA 480  
 DB 421 TCTTCTGAAACATPAACAGACTAGTATTTGATCATTTGATGATGATCTTCTTAA 480  
 QY 481 AGGCGTTTCAATTTTATTTTACAGACGCTTTTATTTAGAGGTCGACATTCATTATGCGGC 540  
 DB 481 AGGCGTTTCAATTTTATTTTACAGACGCTTTTATTTAGAGGTCGACATTCATTATGCGGC 540

QY 541 ATAGGTGTTACATGCGGTATACAACTTAACGGTACACCACTTTTAGCAATGCTCGTAT 600  
 DB 541 ATAGGTGTTACATGCGGTATACAACTTAACGGTACACCACTTTTAGCAATGCTCGTAT 600  
 QY 601 GCGGCACTCTTCCGCTACACAGACCTTTTACATTAATTCTGCTGTTGCAAAACCACT 660  
 DB 601 GCGGCACTCTTCCGCTACACAGACCTTTTACATTAATTCTGCTGTTGCAAAACCACT 660  
 QY 661 GTTAAATATGATCTATGCTGTTCTGCTAGCTTATTTTATTTTAAATGAAAAAAC 720  
 DB 661 GTTAAATATGATCTATGCTGTTCTGCTAGCTTATTTTATTTTAAATGAAAAAAC 720  
 QY 721 ATAAATATGACAAACAAACCTTCCCTTCACTACCTGCTGAGACGACCGCAAGATGGG 780  
 DB 721 ATAAATATGACAAACAAACCTTCCCTTCACTACCTGCTGAGACGACCGCAAGATGGG 780  
 QY 781 TTCAACACGCTGCGACACGATGCAACCGGACCTTCAAGCAATCTGAGGCGCGAC 840  
 DB 781 TTCAACACGCTGCGACACGATGCAACCGGACCTTCAAGCAATCTGAGGCGCGAC 840  
 QY 841 GACGACGTAGGACAGGCGTGCATTAACGACGTGCGCGCATCCAACTTGTCTTCCCTT 900  
 DB 841 GACGACGTAGGACAGGCGTGCATTAACGACGTGCGCGCATCCAACTTGTCTTCCCTT 900  
 QY 901 TCTCTGCTTCACTTCCGCGCGCATCTGCTAGACCCAGGAGATGCTGTGAGAGAGA 960  
 DB 901 TCTCTGCTTCACTTCCGCGCGCATCTGCTAGACCCAGGAGATGCTGTGAGAGAGA 960  
 QY 961 GGTGCGGCGCGCATTTTATATGCTTGGCGGAGACGAGCTTGGCGGAACCGATCCAGA 1020  
 DB 961 GGTGCGGCGCGCATTTTATATGCTTGGCGGAGACGAGCTTGGCGGAACCGATCCAGA 1020  
 QY 1021 GCTCTGCGCAATATCAGCAAGAACCAAGTGGGCGCTGCGCTAGGCCACCGCGCAGAGC 1080  
 DB 1021 GCTCTGCGCAATATCAGCAAGAACCAAGTGGGCGCTGCGCTAGGCCACCGCGCAGAGC 1080  
 QY 1081 GGGGCTTGTTCGAGCGCGTATGCTGCGGAAAGGGAGCGACCCGCTAGGGGGGCCATGCTC 1140  
 DB 1081 GGGGCTTGTTCGAGCGCGTATGCTGCGGAAAGGGAGCGACCCGCTAGGGGGGCCATGCTC 1140  
 QY 1141 CAGGCGCCAGAGAAAAAAGAAAGGAGCGCGGATGATG 1183  
 DB 1141 CAGGCGCCAGAGAAAAAAGAAAGGAGCGCGGATGATG 1183

RESULT 2  
 AAD01016/C  
 ID AAD01016 standard; DNA; 2378 BP.  
 XX  
 AC AAD01016;  
 XX  
 DT 21-SEP-2000 (first entry)  
 XX  
 DE Expression cassette-3 comprising modified E. coli p2A phno coding gene.  
 XX  
 KW Phosphonate herbicide tolerance; aminomethyl phosphonic acid; AMPA; AAT;  
 KW acyltransferase; transacylase; recombinant plant; expression cassette;  
 KW corn; tobacco; wheat; cotton; canola; rice; chloroplast transit peptide;  
 KW CTP; glyphosate oxidase; GOX; glyphosate oxidoreductase; phno gene;  
 KW self-fertilisation; hetero-fertilisation; ds.  
 XX  
 OS Escherichia coli.  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT 28..965  
 FT promoter  
 FT /tag= a  
 FT /note= "Plant functional heterologous promoter"  
 FT 966..1423  
 FT intron  
 FT /tag= b  
 FT /note= "Enhances expression of the phno gene"  
 FT transit\_peptide 1440..1667  
 FT /tag= c



XX Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;  
SQ

Query Match 14.5%; Score 171; DB 3; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 2,4e-74;

Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTATTACCGATATTTT 222  
DB 920 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTATTACCGATATTTT 861  
QY 223 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAAGTAATCTTACCTTTGTTT 282  
DB 860 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAAGTAATCTTACCTTTGTTT 801  
QY 283 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 333  
DB 800 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 750

## RESULT 4

AAC89325/C  
ID AAC89325 standard; DNA; 1501 BP.

AC AAC89325;

DT 07-MAR-2001 (first entry)

DE Rice actin enhancer.

XX 5-enolpyruvylshikimate phosphate synthase; EPSPS; herbicide resistance;  
KM glyphosate; ds.

XX Oryza sp.

OS WO200066747-A1.

XX 09-NOV-2000.

XX 20-APR-2000; 2000WO-GB001572.

XX 29-APR-1999; 99GB-00009967.

XX 29-APR-1999; 99GB-00009969.

XX 29-APR-1999; 99GB-00009972.

XX 29-APR-1999; 99GB-00009981.

XX 29-JUL-1999; 99GB-00017835.

XX 29-JUL-1999; 99GB-00017836.

XX 29-JUL-1999; 99GB-00017843.

XX 21-DEC-1999; 99GB-00030202.

XX 21-DEC-1999; 99GB-00030210.

XX (ZENE) ZENECA LTD.

XX Hawkes TR, Warner SAJ, Andrews CJ, Bachoo S, Pickerill AP;

XX WPI; 2000-679764/66.

XX Isolated polynucleotide encoding a 5-enolpyruvylshikimate phosphate

XX synthase from rice is used for producing transgenic plants with enhanced

XX resistance to glyphosate herbicide.

XX Example 20; Page 54; 98pp; English.

XX The present invention relates to an Oryza sp. 5-enolpyruvylshikimate

XX phosphate synthase (EPSPS) gene. Vectors containing the gene may be used

XX to produce plant tissues and fertile whole plants which are substantially

XX tolerant or substantially resistant to glyphosate herbicide and to

XX produce a herbicidal target which is used for high throughput in vitro

XX screening of potential herbicides

XX Sequence 1501 BP; 495 A; 297 C; 257 G; 452 T; 0 U; 0 Other;

Query Match 14.5%; Score 171; DB 3; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 2,4e-74;

Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTATTACCGATATTTT 222  
DB 920 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTATTACCGATATTTT 861  
QY 223 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAAGTAATCTTACCTTTGTTT 282  
DB 860 ACTTTATACCACTTTTAACTGATGTTTTCACCTTTGACCAAGTAATCTTACCTTTGTTT 801  
QY 283 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 333  
DB 800 TATTTGGACTATCCGACTCTCTTCAAGCATATGATGATGACCTCGAGTA 750

## RESULT 5

ADC84543/C  
ID ADC84543 standard; DNA; 2181 BP.

AC ADC84543;

DT 01-JAN-2004 (first entry)

DE Rice actin promoter encoding sequence.

XX expression cassette; ds.

XX Melanomy caliginosus.

XX WO2003027257-A2.

XX 03-APR-2003.

XX 27-SEP-2002; 2002WO-US030895.

XX 27-SEP-2001; 2001US-0325607P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Lewin JZ, Dietrich R, Budziszewski GJ;

XX WPI; 2003-354651/33.

XX New isolated nucleic acid molecule encoding a polypeptide with a 3'-5'

XX exonuclease domain, useful in molecular biology and transformation, in

XX particular for reproducing and predictably manipulating gene expression

XX in a plant cell.

XX Disclosure; Page 106-107; 108pp; English.

XX The present invention relates to an invention that results in the

XX expression of the endogenous nucleotide sequence in plant cell being

XX increased. In this method the plant cell comprises a second expression

XX cassette comprising a nucleic acid molecule, where the expression of the

XX nucleic acid molecule of interest in the plant cell is decreased as

XX compared to expression of the nucleic acid molecule of interest in a

XX plant cell lacking the first expression cassette. The methods and

XX compositions of the present invention are useful in molecular biology and

XX transformation, in particular for reproducing and predictably

XX manipulating gene expression in a plant cell. The present sequence is a

XX nucleotide sequence of rice actin promoter.

XX Sequence 2181 BP; 602 A; 518 C; 430 G; 631 T; 0 U; 0 Other;

XX Query Match 14.5%; Score 171; DB 10; Length 2181;

XX Best Local Similarity 100.0%; Pred. No. 2,3e-74;

XX Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 163 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTATTACCGATATTTT 222

DB 925 TTTTATATAGTAGAGAAAAGTAATTTACCTTTGGGCGACCTTTATTACCGATATTTT 866



|          |  |   |     |
|----------|--|---|-----|
| Oy       | 223  | ACTTTATACACCTTTTAACTGATGTTTTCACTTTGAACAGGTAATCTTACCTTTGTTT    | 282 |
| Db       | 865  | ACCTTATACACCTTTTAACTGATGTTTTCACTTTGAACAGGTAATCTTACCTTTGTTT    | 806 |
| Oy       | 283  | TATTTTGGACTATCCGACCTCTTCTCTCAAGCATATGATGAATGACTCTGAGTA        | 333 |
| Db       | 805  | TATTTTGGACTATCCGACCTCTTCTCTCAAGCATATGATGAATGACTCTGAGTA        | 755 |
| RESULT 6 |  |   |     |
| ID       | ABK11039/c   | standard; DNA; 9359 BP.                                       |     |
| XX       | ABK11039;  |   |     |
| AC       |  |   |     |
| XX       |  |   |     |
| DT       | 05-JUN-2002  | (first entry)   |     |
| XX       |  |   |     |
| DE       | pVDH636  | vector containing Arabidopsis Ath1 gene.                      |     |
| XX       |  |   |     |
| KW       | pVDH636  | vector; circular; cyclic; grass; plant; herbicide resistance; |     |
| KW       | baseball; cricket; football; golf; rugby; soccer; tennis; lawn; park;    |   |     |
| KW       | athletic field; animal feedstuff; grass flowering inhibitor; Ath1; gene; |   |     |
| XX       | ds.  |   |     |
| XX       |  |   |     |
| OS       | Arabidopsis thaliana.  |   |     |
| OS       | Zea mays.  |   |     |
| OS       | Agrobacterium tumefaciens.   |   |     |
| OS       | Oryza sativa.  |   |     |
| OS       | Escherichia coli.  |   |     |
| OS       | Cauliflower mosaic virus.  |   |     |
| OS       | Synthetic.   |   |     |
| XX       |  |   |     |
| FH       | Key  | Location/Qualifiers   |     |
| FT       | CDS  | 4930..6360  |     |
| FT       |  | /*tag= a  |     |
| FT       |  | /product= "Ath1"  |     |
| FT       |  | /partial  |     |
| FT       |  | /note= "No stop codon"  |     |
| XX       |  |   |     |
| PN       | WO200214524-A2.  |   |     |
| PD       |  |   |     |
| XX       | 21-FEB-2002.   |   |     |
| PF       | 16-AUG-2001; 2001WO-EP009572.  |   |     |
| XX       |  |   |     |
| PR       | 18-AUG-2000; 2000US-0226422P.  |   |     |
| PR       | 27-NOV-2000; 2000US-0253274P.  |   |     |
| PR       | 27-NOV-2000; 2000US-0253337P.  |   |     |
| PR       | 22-JUN-2001; 2001US-0300220P.  |   |     |
| XX       |  |   |     |
| PA       | (ADVA-) ADVANTA SEEDS BV.  |   |     |
| PI       | Van Der Valk P, Van Dun CMP, Smeekens SCM, Proveniers MCG;               |   |     |
| XX       |  |   |     |
| DR       | WPI: 2002-257611/30.   |   |     |
| DR       | P-PSDB; AAU76514.  |   |     |
| XX       |  |   |     |
| PT       | New genetically modified grasses that express inhibited generative       |   |     |
| PT       | propagation, or herbicide resistance, useful for forage (e.g. cattle     |   |     |
| PT       | feedstuff) or amenity purposes (e.g. for use in an athletic field, lawn  |   |     |
| PT       | or park).  |   |     |
| XX       |  |   |     |
| PS       | Example 1; Fig 6; 56pp; English.   |   |     |
| XX       |  |   |     |
| CC       | The invention relates to a grass plant, which has been genetically       |   |     |
| CC       | modified to substantially inhibit generative propagation and carry       |   |     |
| CC       | herbicide resistance. The grass is useful for growth and/or propagation  |   |     |
| CC       | of grasses. The grass is used to play at least one sport (e.g. baseball, |   |     |
| CC       | cricket, football, golf, rugby, soccer or tennis), or used at least in a |   |     |
| CC       | portion of an athletic field, lawn or park. The grass is fed to animal   |   |     |
| CC       | (e.g. cattle, goat, horse or sheep) or used as an animal feedstuff. The  |   |     |
| CC       | present sequence represents the coding sequence of pVDH636 vector        |   |     |

| CC | containing the Arabidopsis Ach1 gene, which was used to inhibit flowering in grasses |
|----|--|
| CC | Sequence 9359 BP; 2377 A; 2222 C; 2133 G; 2624 T; 0 U; 3 Other;                      |
| CC | Query Match 14.5%; Score 171; DB 6; Length 9359;                                     |
| CC | Best Local Similarity 100.0%; Pred. No. 2.2e-74;                                     |
| CC | Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;                         |
| CC | 163 TTTTATATAGTAGAAAAGAGTAATTTCACTTTGGCCACCTTTTATTAACGATATTT 222                     |
| CC | DB 6658 TTTTATATAGTAGAAAAGAGTAATTTCACTTTGGCCACCTTTTATTAACGATATTT 6799                |
| CC | DB 223 ACTTATATACCACTTTTACTGATGTTTTCACCTTTTGACCGAGTAATCTTAACCTTTGTT 282              |
| CC | DB 6798 ACTTATATACCACTTTTAACTGAGTGTTCACCTTTTGACCGAGTAATCTTAACCTTTGTT 6799            |
| CC | QY 283 TATTTTGACATCCCGACTCTCTTCTCAAGCATATGATGATGACCTCGAGTA 333                       |
| CC | DB 6738 TATTTTGACATCCCGACTCTCTTCTCAAGCATATGATGATGACCTCGAGTA 6688                     |
| CC | RESULT 7   |
| CC | ABK10687/C   |
| CC | ID ABK10687 standard; DNA; 9359 BP.  |
| CC | AC ABK10687;   |
| CC | XX   |
| CC | DT 05-JUN-2002 (first entry)   |
| CC | XX   |
| CC | DE Transformation vector plasmid pVDH636 DNA.  |
| CC | XX   |
| CC | KW Grass; plant; transgenic; flowering inhibition; inflorescence; gene; de;          |
| CC | KW tiller production; delayed heading; gibberellic acid; phytohormone; golf;         |
| CC | KW genetically modified grass; athletic field; sport; baseball; cricket;             |
| CC | KW football; rugby; soccer; tennis; lawn; landscaping; cattle; horse; sheep;         |
| CC | KW goat; animal feedstuff; Arabidopsis thaliana; Ath1; homeobox gene;                |
| CC | KW circular; cyclic; pVDH636.  |
| CC | XX   |
| CC | OS Synthetic.  |
| CC | XX   |
| CC | FH Key Location/Qualifiers   |
| CC | FT CDS 4930..6351  |
| CC | FT /*tag= a  |
| CC | FT /product= "Arabidopsis thaliana Ach1"   |
| CC | XX   |
| CC | PN WO200214486-A2.   |
| CC | XX   |
| CC | PD 21-FEB-2002.  |
| CC | XX   |
| CC | PF 16-AUG-2001; 2001WO-EP009570.   |
| CC | XX   |
| CC | PR 18-AUG-2000; 2000US-0226422P.   |
| CC | PR 27-NOV-2000; 2000US-0253274P.   |
| CC | PR 27-NOV-2000; 2000US-0253327P.   |
| CC | PR 22-JUN-2001; 2001US-0300220P.   |
| CC | XX   |
| CC | PA (ADVA-) ADVANTA SEEDS BV.   |
| CC | XX   |
| CC | PI Van Der Valk P, Van Dun CMP, Smeekens SCM, Proveniers MCG;                        |
| CC | XX   |
| CC | DR WPI; 2002-257603/30.  |
| CC | DR P-PSDB; AAU76885.   |
| CC | XX   |
| CC | PT New genetically modified grass useful for growing and/or propagating              |
| CC | PT grass in athletic fields (for sports such as baseball, cricket, football,         |
| CC | PT golf, rugby, soccer and tennis), or as animal feedstuff for cattle, goat,         |
| CC | PT horse and sheep.  |
| CC | XX   |
| CC | PS Example 1; Fig 6; 45pp; English.  |
| CC | XX   |
| CC | XX The invention relates to a grass plant which has been genetically                 |
| CC | modified to substantially inhibit generative propagation. The genetic                |



```

XX Synthetic.
OS
XX Key
FH misc_feature
FT 1. .1146
FT /tag= a
FT /note= "tobacco Rb7 MARs"
FT
FT misc_feature
FT 1167. .1304
FT /tag= b
FT /note= "linker sequence (ACF58351)"
FT
FT promoter
FT 1305. .2701
FT /tag= c
FT /note= "rice actin promoter and intron"
FT
FT intron
FT 2235. .2696
FT /tag= d
FT /note= "rice actin intron"
FT
FT misc_feature
FT 2704. .3258
FT /tag= e
FT /note= "PAT gene"
FT
FT misc_feature
FT 3259. .3272
FT /tag= f
FT /note= "linker sequence (ACF58352)"
FT
FT 3'UTR
FT 3273. .3629
FT /tag= g
FT /note= "maize lipase UTR"
FT
FT misc_feature
FT complement (3671. .4836)
FT /tag= i
FT /note= "tobacco Rb7 MAR"
FT
FT misc_feature
FT 4351. .4405
FT /tag= h
FT /note= "linker sequence (ACF58353)"
FT
FT misc_feature
FT 4837. .4857
FT /tag= j
FT /note= "linker sequence (ACF58354)"
FT
FT misc_feature
FT 4858. .5103
FT /tag= k
FT /note= "Puc19"
FT
FT misc_feature
FT 5104. .5130
FT /tag= l
FT /note= "linker sequence (ACF58355)"
FT
FT misc_feature
FT 5130. .7523
FT /tag= m
FT /note= "Puc19"
FT
FT misc_feature
FT 7524. .7545
FT /tag= n
FT /note= "linker sequence (ACF58356)"
XX
XX WO2003078614-A2.
XX
XX 25-SEP-2003.
XX
XX 18-MAR-2003; 2003WO-IB001562.
XX
XX 19-MAR-2002; 2002US-0365769P.
XX
XX 26-MAR-2002; 2002US-0368047P.
XX
XX (PLAN-) PLANT RES INT BV.
XX
XX Bakker HAC, Florack DEA, Bosch HJ;
XX
XX WPI; 2003-779132/73.
XX
XX New plant host cell system for producing a desired glycoprotein comprises
XX a mammalian N-acetylglucosaminyltransferase (GntIII) enzyme, a nucleic
XX acid sequence encoding the enzyme, or a vector comprising the nucleic
XX acid sequence.
XX
XX Example 6; Fig 7B; 122pp; English.
XX
XX The invention relates to a plant host cell system comprising a mammalian
XX UDP-N-acetylglucosamine: beta-D-mannoside beta(1,4)-N-
XX acetylglucosaminyltransferase (GntIII) enzyme, a nucleic acid sequence
XX encoding a mammalian GntIII protein, or a vector comprising the GntIII

```

```

CC nucleic acid. The plant host system is useful in producing a desired
CC glycoprotein or its functional fragment. The glycoprotein or its
CC functional fragment may be used for the production of a pharmaceutical
CC composition. The present sequence represents the nucleotide sequence of
CC plasmid pDA8504
XX
XX SO Sequence 7545 BP; 2128 A; 1579 C; 1622 G; 2216 T; 0 U; 0 Other;
XX
XX Query Match 14.3%; Score 169; DB 10; Length 7545;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-73;
XX Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 163 TTTTAAATAGTGAAGAAGTAATTTCACTTTGGGCCACTTTATTATCCGATATTT 222
XX |||||
XX DB 1473 TTTTAAATAGTGAAGAAGTAATTTCACTTTGGGCCACTTTATTATCCGATATTT 1414
XX |||||
XX QY 223 ACTTTAACCACCTTTTAACTGATGTTTCACTTTGACCAAGTAATCTTACCTTTGTT 282
XX |||||
XX DB 1413 ACTTATACCACTTTTAACTGATGTTTCACTTTGACCAAGTAATCTTACCTTTGTT 1354
XX |||||
XX QY 283 TATTTGACTATCCGACTCTTCTCAAGCATATGATGACTGAG 331
XX |||||
XX DB 1353 TATTTGACTATCCGACTCTTCTCAAGCATATGATGACTGAG 1305
XX |||||
XX
XX RESULT 10
XX ACF58344/C
XX ID ACF58344 standard; DNA; 11643 BP.
XX
XX AC ACF58344;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX XX Nucleotide sequence of plasmid pDA87113.
XX DE
XX XX GntIII; N-acetylglucosaminyltransferase; transgenic; glycoprotein; ds.
XX KM
XX XX Synthetic.
XX OS
XX FH Key
XX FH Location/Qualifiers
XX FT 1. .1164
XX FT /tag= a
XX FT /note= "Rb7 MAR v3"
XX FT 1165. .1233
XX FT /tag= b
XX FT /note= "linker sequence (ACF58357)"
XX FT 1234. .3224
XX FT /tag= c
XX FT /note= "maize ubiquitin 1 promoter"
XX FT 3225. .4891
XX FT /tag= d
XX FT /note= "GntIII v.2"
XX FT 4896. .5260
XX FT /tag= e
XX FT /note= "maize peroxidase-5 3'-UTR"
XX FT 5261. .5404
XX FT /tag= f
XX FT /note= "multiple cloning sites (ACF58358)"
XX FT 5405. .6802
XX FT /tag= g
XX FT /note= "rice actin 1 promoter v2"
XX FT 6803. .7358
XX FT /tag= h
XX FT /note= "PAT v3"
XX FT 7359. .7372
XX FT /tag= i
XX FT /note= "linker sequence (ACF58359)"
XX FT 7373. .7729
XX FT /tag= j
XX FT /note= "maize lipase UTR"
XX FT 7730. .7770
XX FT /tag= k
XX FT /note= "linker sequence (ACF58360)"

```

```

FT misc_feature 7771..8934
FT /*tag= 1
FT /note= "Rp7 MAR v3"
FT misc_feature 8935..11643
FT /*tag= m
FT /note= "Puc19"
FT misc_feature 10164..11021
FT /*tag= n
FT /note= "ampicillin resistance gene"
XX PN MO2003078614-A2.
XX PD 25-SEP-2003.
XX PF 18-MAR-2003; 2003WO-1B001562.
XX PR 19-MAR-2002; 2002US-0365769P.
XX PR 26-MAR-2002; 2002US-0368047P.
XX PA (PLAN-) PLANT RES INT BV.
XX PI Bakker HAC, Florack DEA, Bosch HJ;
XX DR WPI; 2003-779132/73.
XX PS Example 6; Fig 8B; 122pp; English.
XX CC The invention relates to a plant host cell system comprising a mammalian
XX CC UDP-N-acetylglucosamine: beta-D-mannoside beta(1,4)-N-
XX CC acetylglucosaminyltransferase (GntII) enzyme, a nucleic acid sequence
XX CC encoding a mammalian GntII protein, or a vector comprising the GntII
XX CC nucleic acid. The plant host system is useful in producing a desired
XX CC glycoprotein or its functional fragment. The glycoprotein or its
XX CC functional fragment may be used for the production of a pharmaceutical
XX CC composition. The present sequence represents the nucleotide sequence of
XX CC plasmid pDA87113
XX SQ Sequence 11643 BP; 3005 A; 2672 C; 2635 G; 3331 T; 0 U; 0 Other;
XX
XX Query Match 14.3%; Score 169; DB 10; Length 11643;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-73;
XX Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 163 TTTTATATAGTAGAGAAAGAGTAATTTCACTTTGGGCCACCTTTTATTACCGATATTTT 222
XX DB 5573 TTTTATATAGTAGAGAAAGAGTAATTTCACTTTGGGCCACCTTTTATTACCGATATTTT 514
XX
XX QY 223 ACTTATACCACTTTTAACTGATGTTTTCCTTTTGACCAAGTAATTTTACCTTTGTTT 282
XX DB 5513 ACTTATACCACTTTTAACTGATGTTTTCCTTTTGACCAAGTAATTTTACCTTTGTTT 5454
XX
XX QY 283 TATTTTGACATATCCGACTCTCTTCTCAAGCATATGATGACCTCGAG 331
XX DB 5453 TATTTTGACATATCCGACTCTCTTCTCAAGCATATGATGACCTCGAG 5405
XX
XX RESULT 11
XX AAV23237/c
XX ID AAV23237 standard; DNA; 4032 BP.
XX AC AAV23237;
XX AC 17-JUL-1998 (first entry)
XX DT
XX XX Plasmid pMV71.
XX DE
XX XX Barstar; barnase inhibitor; fertility restoration; male-sterile line;
XX KW plasmid pMV71; circular; ds.

```

```

XX OS Synthetic.
XX XX Location/Qualifiers
XX FH Key 1999..3400
XX FT promoter
XX FT /*tag= a
XX FT /note= "promoter region of rice actin gene - contains an
XX FT intron in the leader"
XX FT CDS 3401..3676
XX FT /*tag= b
XX FT /product= "barstar"
XX FT 3'UTR 3677..4003
XX FT /*tag= c
XX FT /note= "region containing 3' untranslated end of the
XX FT nopaline synthase gene of Agrobacterium T-DNA"
XX
XX PN MO9810081-A2.
XX PD 12-MAR-1998.
XX PF 01-SEP-1997; 97WO-EP004739.
XX PF 03-SEP-1996; 96EP-00202446.
XX PR (PUBZ ) PLANT GENETIC SYSTEMS NV.
XX PA
XX PI Michèle F, Williams M;
XX DR WPI; 1998-193630/17.
XX PS Example 3; Page 37-39; 54pp; English.
XX CC The present sequence was used in the preparation of an improved Bacillus
XX CC amyloliquefaciens barstar, i.e. barnase inhibitor, which can be used to
XX CC restore fertility to male-sterile lines. The DNA sequence encoding the
XX CC improved barstar, leads to increased barstar production in tapetum cells,
XX CC due to improved translation, and possibly protein stability
XX CC
XX SQ Sequence 4032 BP; 1072 A; 968 C; 963 G; 1029 T; 0 U; 0 Other;
XX
XX Query Match 14.2%; Score 168; DB 2; Length 4032;
XX Best Local Similarity 100.0%; Pred. No. 7e-73;
XX Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 163 TTTTATATAGTAGAGAAAGAGTAATTTCACTTTGGGCCACCTTTTATTACCGATATTTT 222
XX DB 2166 TTTTATATAGTAGAGAAAGAGTAATTTCACTTTGGGCCACCTTTTATTACCGATATTTT 2107
XX
XX QY 223 ACTTATACCACTTTTAACTGATGTTTTCCTTTTGACCAAGTAATTTTACCTTTGTTT 282
XX DB 2106 ACTTATACCACTTTTAACTGATGTTTTCCTTTTGACCAAGTAATTTTACCTTTGTTT 2047
XX
XX QY 283 TATTTTGACATATCCGACTCTCTTCTCAAGCATATGATGACCTCGA 330
XX DB 2046 TATTTTGACATATCCGACTCTCTTCTCAAGCATATGATGACCTCGA 1999
XX
XX RESULT 12
XX AAF80287
XX ID AAF80287 standard; DNA; 6865 BP.
XX AC AAF80287;
XX AC 29-JUN-2001 (first entry)
XX DT
XX XX Nucleotide sequence of plasmid pMR1195.
XX DE
XX XX Vector; transgene(s); trfA locus; RK2 ori; oriV; P285 protein;
XX KW P382 protein; antibiotic resistance gene; nptII; transgenic plant; ss.
XX

```

| OS | Synthetic.   | Location/Qualifiers   |
|----|--|---|
| XX | Key  | 1..654  |
| FH | rep_origin   | /tag= a   |
| FT |  | /note= "ori RK2"  |
| FT |  | 655..1263   |
| FT | rep_origin   | /tag= b   |
| FT |  | /note= "ori COLE1"  |
| FT |  | 1264..2603  |
| FT | CDS  | /tag= c   |
| FT |  | /note= "NPT III gene coding for neomycin phosphotransferase and kanamycin resistance"                           |
| FT |  | 2604..4098  |
| FT | CDS  | /tag= d   |
| FT |  | /note= "trfA locus from RK2 coding for two proteins P285 and P382 enabling an increase in the replication rate" |
| FT |  | 4106..4271  |
| FT | misc_feature   | /tag= e   |
| FT |  | /note= "T-DNA left border"  |
| FT |  | 4272..4559  |
| FT | terminator   | /tag= f   |
| FT |  | /note= "nopaline synthetase terminator"   |
| FT |  | 4575..5150  |
| FT | CDS  | /tag= g   |
| FT |  | /note= "Bar gene coding for phosphinotricin acetyltransferase and glufosinate resistance"                       |
| FT |  | 5174..5685  |
| FT | intron   | /tag= h   |
| FT |  | /note= "rice actin intron"  |
| FT |  | 5686..6626  |
| FT | promoter   | /tag= i   |
| FT |  | /note= "rice actin promoter"  |
| FT |  | 6685..6858  |
| FT | misc_feature   | /tag= j   |
| FT |  | /note= "T-DNA right border"   |
| FT |  |   |
| PN | FR2798139-A1.  |   |
| XX | 09-MAR-2001.   |   |
| XX | 03-SEP-1999;   | 99FR-00011112.  |
| XX | 03-SEP-1999;   | 99FR-00011112.  |
| XX | (MER1-) MERISTEM THERAPEUTICS SA.  |   |
| XX | Gruber V, Comeau D;  |   |
| XX | WPI; 2001-259847/27.   |   |
| DR | New vector free from non-essential elements, useful for transforming cells for protein production and for preparing transgenic plants.   |   |
| PT | Claim 20; Page 128-131; 180pp; French.   |   |
| PS | The specification describes a synthetic vector containing only those elements essential for its functionality and transgenesis of a cell (especially a plant cell). The vector consists of at most one origin of replication (ori), at most one sequence encoding a selection agent and a trfA locus encoding a protein that increases the level of plasmid replication. The vector particularly contains an RK2 ori, especially oriIy from PRK2 of Escherichia coli with a broad host range, an antibiotic resistance gene (especially nptIII conferring resistance to kanamycin in bacteria) and a trfA locus from PRK2 encoding the proteins P285 and P382. The vectors are used to prepare transgenic plants and transformed host cells for production of a heterologous proteins, e.g. insulin, interferon, lipase, blood proteins and anti-inflammatory agents. The present sequence represents a plasmid of the invention |   |
| CC | Sequence 6865 BP; 1699 A; 1694 C; 1919 G; 1553 T; 0 U; 0 Other;  |   |
| XX | Query Match 13.8%; Score 163; DB 4; Length 6865;   |   |

|                  | Best Local Similarity   | 100.0%; | Pred. No.2.le-70; | Mismatches | 0;     | Indels | 0;   | Gaps | 0;   |
|------------------|---|---------|-------------------|------------|--------|--------|------|------|------|
| Oy               | Matches 163; Conservative   | 0;      | Mismatches        | 0;         | Indels | 0;     | Gaps | 0;   |      |
| Dd               | 163<br>TTTTATATAGTGAAGAAGATAATTTCACCTTTGGGCGACCTTTTATTAACCATATTTT   |         |                   |            |        |        |      |      | 222  |
| Oy               | 6383<br>TTTTATATATAGTAGAAGAAGATAATTTCACCTTTGGGCGACCTTTTATTAACCATATTTT   |         |                   |            |        |        |      |      | 6442 |
| Dd               | 223<br>ACTTTATACCACTTTTAACTGATGTTTTGACCTTTTGACCAGATATCTTAACCTTTGTTT   |         |                   |            |        |        |      |      | 282  |
| Oy               | 6443<br>ACTTTATACCACTTTTAACTGATGTTTTGACCTTTTGACCAGATATCTTAACCTTTGTTT  |         |                   |            |        |        |      |      | 6502 |
| Dd               | 283<br>TATTTTGGACTATCCGCACCTCTCTTTCGACGATATGAATGAC  |         |                   |            |        |        |      |      | 325  |
| Oy               | 6503<br>TATTTTGGACTATCCGCACCTCTCTTTCGACGATATGAATGAC   |         |                   |            |        |        |      |      | 6545 |
| <b>RESULT_13</b> |   |         |                   |            |        |        |      |      |      |
| AF80295          |   |         |                   |            |        |        |      |      |      |
| ID               | AAF80295 standard; DNA; 10003 BP.   |         |                   |            |        |        |      |      |      |
| XX               |   |         |                   |            |        |        |      |      |      |
| AC               | AAF80295;   |         |                   |            |        |        |      |      |      |
| XX               |   |         |                   |            |        |        |      |      |      |
| DT               | 29-JUN-2001 (first entry)   |         |                   |            |        |        |      |      |      |
| XX               |   |         |                   |            |        |        |      |      |      |
| DE               | Nucleotide sequence of plasmid pMR1210.   |         |                   |            |        |        |      |      |      |
| XX               |   |         |                   |            |        |        |      |      |      |
| XX               | Vector; transgenesis; trfA locus; RK2 ori; oriV; P285 protein;  |         |                   |            |        |        |      |      |      |
| KW               | P382 protein; antibiotic resistance gene; nptIII; transgenic plant; ss.   |         |                   |            |        |        |      |      |      |
| XX               |   |         |                   |            |        |        |      |      |      |
| OS               | Synthetic.  |         |                   |            |        |        |      |      |      |
| XX               |   |         |                   |            |        |        |      |      |      |
| FH               | Key   |         |                   |            |        |        |      |      |      |
| FT               | rep_origin  |         |                   |            |        |        |      |      |      |
| FT               |   |         |                   |            |        |        |      |      |      |
| FT               | Location/Qualifiers   |         |                   |            |        |        |      |      |      |
| FT               | 1..654  |         |                   |            |        |        |      |      |      |
| FT               | /*tag= a  |         |                   |            |        |        |      |      |      |
| FT               | /note= "ori RK2"  |         |                   |            |        |        |      |      |      |
| FT               | 655..1263   |         |                   |            |        |        |      |      |      |
| FT               | /*tag= b  |         |                   |            |        |        |      |      |      |
| FT               | /note= "ori ColE1"  |         |                   |            |        |        |      |      |      |
| FT               | 1264..2603  |         |                   |            |        |        |      |      |      |
| FT               | /*tag= c  |         |                   |            |        |        |      |      |      |
| FT               | /note= "NPT III gene coding for neomycin phosphotransferase and kanamycin resistance"                           |         |                   |            |        |        |      |      |      |
| FT               | 2604..4098  |         |                   |            |        |        |      |      |      |
| FT               | /*tag= d  |         |                   |            |        |        |      |      |      |
| FT               | /note= "trfA locus from RK2 coding for two proteins P285 and P382 enabling an increase in the replication rate" |         |                   |            |        |        |      |      |      |
| FT               | 4106..4271  |         |                   |            |        |        |      |      |      |
| FT               | /*tag= e  |         |                   |            |        |        |      |      |      |
| FT               | /note= "T-DNA left border"  |         |                   |            |        |        |      |      |      |
| FT               | 4272..4559  |         |                   |            |        |        |      |      |      |

/note="nopaline synthetase terminator"  
9823. .9996  
/\*tag= n  
/note="T-DNA right border"

FR2798139-A1.  
XX  
XX 09-MAR-2001.  
XX  
XX 03-SEP-1999; 99FR-00011112.  
XX  
XX 03-SEP-1999; 99FR-00011112.  
XX  
XX (MERI-) MERISTEM THERAPEUTICS SA.  
XX  
XX Gruber V, Comeau D;  
XX  
XX WPI; 2001-259847/27.  
XX  
XX  
XX New vector free from non-essential elements, useful for transforming  
XX cells for protein production and for preparing transgenic plants.  
XX  
XX Claim 20; Page 163-166; 180pp; French.

The specification describes a synthetic vector containing only those  
XX elements essential for its functionality and transgenesis of a cell  
XX (especially a plant cell). The vector consists of at most one origin of  
XX replication (ori), at most one sequence encoding a selection agent and a  
XX trfA locus encoding a protein that increases the level of plasmid  
XX replication. The vector particularly contains an RK2 ori, especially oriV  
XX from PRK2 of Escherichia coli with a broad host range, an antibiotic  
XX resistance gene (especially nptII conferring resistance to kanamycin in  
XX bacteria) and a trfA locus from PRK2 encoding the proteins P285 and P382.  
XX The vectors are used to prepare transgenic plants and transformed host  
XX cells for production of a heterologous proteins, e.g. insulin,  
XX interferon, lipase, blood proteins and anti-inflammatory agents. The  
XX present sequence represents a plasmid of the invention

Sequence 10003 BP; 2476 A; 2470 C; 2729 G; 2328 T; 0 U; 0 Other;  
SQ

Query Match 13.8%; Score 163; DB 4; Length 10003;  
Best Local Similarity 100.0%; Pred. No. 2.1e-70;  
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAGAAAGAGTAATTCACCTTGGGCGACCTTTATACGATATTT 222  
DB TTTTATATAGTAGAGAAAGAGTAATTCACCTTGGGCGACCTTTATACGATATTT 6461  
QY 223 ACTTATACGACCTTTTAACTGATGTTTCACTTTTGACCAAGTAATCTTACCTTGT 282  
DB ACTTATACGACCTTTTAACTGATGTTTCACTTTTGACCAAGTAATCTTACCTTGT 6521  
QY 283 TATTTGGACTATCCCGACTCTCTTCTCAGCATATGATGAC 325  
DB TATTTGGACTATCCCGACTCTCTTCTCAGCATATGATGAC 6564  
6522 TATTTGGACTATCCCGACTCTCTTCTCAGCATATGATGAC 6564

RESULT 14  
ADP83768 standard; DNA; 474 BP.  
XX  
XX ADP83768;  
XX  
XX 09-SEP-2004 (first entry)  
XX  
XX Bentgrass event ASR-368 3' transgene/genomic chimeric DNA SEQ ID NO:4.  
XX  
XX ds: bentgrass: ASR-368; turfgrass stand; herbicide tolerance;  
XX glyphosate herbicide.  
XX Agrobacterium tumefaciens.  
XX Chimeric.  
XX Unidentified.

XX  
XX W02004053062-A2.  
XX  
XX 24-JUN-2004.  
XX  
XX 03-DEC-2003; 2003WO-US038268.  
XX  
XX 05-DEC-2002; 2002US-0431153P.  
XX  
XX (MONS ) MONSANTO TECHNOLOGY LLC.  
XX (SCOT-) SCOTTS CO.  
XX  
XX Guo SX, Harriman RW, Lee L, Nelson EK;  
XX  
XX WPI; 2004-468838/44.  
XX  
XX  
XX New bentgrass ASR-368 plant and seed useful for producing plants with  
XX improved glyphosate herbicide tolerance or as a turfgrass stand.  
XX  
XX Claim 5; SEQ ID NO 4; 37pp; English.

The invention relates to a novel seed of the bentgrass plant designated  
XX ASR-368. Also disclosed is the method for controlling weeds in a  
XX turfgrass stand of bentgrass ASR-368. The plant or plant parts are useful  
XX for producing plants with improved agronomic traits, such as herbicide  
XX tolerance, particularly to glyphosate herbicide. It may be used as a  
XX turfgrass stand. Methods are disclosed that are used for detecting the  
XX transgene/genomic junction region in bentgrass plant ASR-368 or its  
XX progeny. The present sequence represents the ASR-368 3' transgene/genomic  
XX DNA.

Sequence 474 BP; 111 A; 144 C; 110 G; 109 T; 0 U; 0 Other;  
SQ

Query Match 13.4%; Score 159; DB 12; Length 474;  
Best Local Similarity 100.0%; Pred. No. 2.3e-68;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGGTTTATGATTAAGTCCGCAATTATACATTAAATACCG 60  
DB GACGTTATTTATGAGATGGGTTTATGATTAAGTCCGCAATTATACATTAAATACCG 149  
QY 61 GATGAAACAAATATATGAGGCGCAACTAGTAATTTTCCGGCGGCGTGCATCTAT 120  
DB GATGAAACAAATATATGAGGCGCAACTAGTAATTTTCCGGCGGCGTGCATCTAT 209  
QY 121 GTTACTAGATGGGATATCCCGGGGGAATTCGGTACCA 159  
DB GTTACTAGATGGGATATCCCGGGGGAATTCGGTACCA 248  
210 GTTACTAGATGGGATATCCCGGGGGAATTCGGTACCA 248

RESULT 15  
AAD51491 standard; DNA; 349 BP.  
XX  
XX AAD51491;  
XX  
XX 16-APR-2003 (first entry)  
XX  
XX A. tumefaciens DNA used in the exemplification of the invention.  
XX  
XX Insect resistance; MON15985 event; plant breeding; ds.  
XX Agrobacterium tumefaciens.  
XX  
XX W02002100163-A2.  
XX  
XX 19-DEC-2002.  
XX  
XX 05-JUN-2002; 2002WO-US017853.  
XX  
XX 11-JUN-2001; 2001US-0297406P.  
XX  
XX (MONS ) MONSANTO TECHNOLOGY LLC.



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic acid search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 1667.4 Seconds  
(without alignments)  
3460.943 Million cell updates/sec

Title: US-10-790-430-8

Perfect score: 1183

Sequence: 1 gacgtatctatcagatgagtg...aaggaagcgcgagatgatg 1183

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                     | Description       |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1          | 1183  | 100.0       | 1183   | 9 US-09-872-051-8      | Sequence 8, Appli |
| 2          | 1183  | 100.0       | 1183   | 18 US-10-790-430-8     | Sequence 8, Appli |
| 3          | 305   | 25.8        | 6683   | 18 US-10-425-115-58081 | Sequence 58081, A |
| 4          | 202   | 17.1        | 2378   | 14 US-10-213-791-27    | Sequence 27, Appl |
| 5          | 171   | 14.5        | 1501   | 14 US-10-012-070A-49   | Sequence 49, Appl |
| 6          | 171   | 14.5        | 1501   | 14 US-10-012-013-44    | Sequence 44, Appl |
| 7          | 169   | 14.3        | 2480   | 19 US-10-678-588A-1    | Sequence 1, Appli |
| 8          | 168   | 14.2        | 4032   | 9 US-09-970-921-5      | Sequence 5, Appli |
| 9          | 163   | 13.8        | 6865   | 10 US-09-845-064-13    | Sequence 11, Appl |
| 10         | 163   | 13.8        | 10003  | 10 US-09-845-064-21    | Sequence 21, Appl |
| 11         | 147   | 12.4        | 349    | 18 US-10-480-223A-23   | Sequence 23, Appl |

12 147 12.4 1360 18 US-10-480-223A-12 Sequence 12, Appl  
13 147 12.4 3754 15 US-10-233-665-15 Sequence 15, Appl  
14 147 12.4 4149 15 US-10-233-665-13 Sequence 13, Appl  
15 147 12.4 8349 16 US-10-198-478-16 Sequence 16, Appl  
16 147 12.4 10249 16 US-10-198-478-14 Sequence 14, Appl  
17 147 12.4 10312 16 US-10-198-478-15 Sequence 15, Appl  
18 147 12.4 10339 16 US-10-198-478-13 Sequence 13, Appl  
19 144 12.2 11546 18 US-10-841-796-33 Sequence 27, Appl  
20 140 11.8 2378 15 US-10-213-791-27 Sequence 31, Appl  
21 140 11.8 2436 15 US-10-213-791-21 Sequence 23, Appl  
22 137 11.6 1630 15 US-10-213-791-23 Sequence 9, Appl  
23 137 11.6 6313 13 US-10-047-542-9 Sequence 14, Appl  
24 137 11.6 8074 13 US-10-047-542-14 Sequence 5, Appl  
25 137 11.6 10846 9 US-09-923-109-5 Sequence 5, Appl  
26 137 11.6 10846 15 US-10-164-204-5 Sequence 5, Appl  
27 137 11.6 10846 17 US-10-705-430-5 Sequence 6, Appl  
28 137 11.6 10900 9 US-09-923-109-6 Sequence 6, Appl  
29 137 11.6 10900 15 US-10-164-204-6 Sequence 6, Appl  
30 137 11.6 10900 17 US-10-705-430-6 Sequence 6, Appl  
31 137 11.6 11606 18 US-10-602-475A-9 Sequence 69, Appl  
32 136 11.5 1681 18 US-10-168-653C-69 Sequence 71, Appl  
33 136 11.5 1851 18 US-10-168-653C-71 Sequence 70, Appl  
34 136 11.5 2315 18 US-10-168-653C-70 Sequence 67, Appl  
35 136 11.5 2708 18 US-10-168-653C-67 Sequence 67, Appl  
36 136 11.5 11522 14 US-10-052-092-19 Sequence 19, Appl  
37 136 11.5 11522 16 US-10-437-107-19 Sequence 19, Appl  
38 136 11.5 11522 18 US-10-896-419-19 Sequence 104, App  
39 134 11.3 277 15 US-10-161-403-104 Sequence 107, App  
40 134 11.3 277 15 US-10-161-403-107 Sequence 15, Appl  
41 134 11.3 277 18 US-10-161-408-15 Sequence 18, Appl  
42 134 11.3 277 18 US-10-161-408-18 Sequence 6, Appl  
43 134 11.3 563 9 US-09-970-921-6 Sequence 16, Appl  
44 134 11.3 1213 9 US-09-794-384A-16 Sequence 16, Appl  
45 134 11.3 1213 18 US-10-801-550-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-09-872-051-8  
Sequence 8, Application US/09872051  
Patent No. US20020013960A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(Ink603) and Composition and Methods for D  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 1183  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(1183)  
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
OTHER INFORMATION: 165-381 construct vector DNA  
OTHER INFORMATION: 382-686 Zea maize plastid gene, rps11 and rpoA  
OTHER INFORMATION: 687-1183 Zea maize genomic DNA

US-09-872-051-8

Query Match 100.0%; Score 1183; DB 9; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATAGCG 60  
DB 1 GACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATAGCG 60  
QY 61 GATGAAAAAATAATATAGCGCGCAAACTAGATATTAATTCGCGCGGCTCATCTAT 120  
DB 61 GATGAAAAAATAATATAGCGCGCAAACTAGATATTAATTCGCGCGGCTCATCTAT 120  
QY 121 GTTACTAGATCGGGGATATCCCGGGGAATTCGATACCAAGCTTTATATAGTAAAA 180  
DB 121 GTTACTAGATCGGGGATATCCCGGGGAATTCGATACCAAGCTTTATATAGTAAAA 180  
QY 181 GAGTAAATTTACATTTGGGCACTTTTATACGATATTTTACTTTATACACCTTTTA 240  
DB 181 GAGTAAATTTACATTTGGGCACTTTTATACGATATTTTACTTTATACACCTTTTA 240  
QY 241 ACTGATGTTTACATTTTGACCAAGTATCTTACTTTGTTTATTTTGACATATCCGA 300  
DB 241 ACTGATGTTTACATTTTGACCAAGTATCTTACTTTGTTTATTTTGACATATCCGA 300  
QY 301 CTCTCTTCTCAAGCATATGATGACCTCGAGTAACTGTTAAACGCGCGCCCTAGGGA 360  
DB 301 CTCTCTTCTCAAGCATATGATGACCTCGAGTAACTGTTAAACGCGCGCCCTAGGGA 360  
QY 361 TATCAAGCTTGTATGACCAAGCACTTCTGATGTTTGAAGTATCTGTTATC 420  
DB 361 TATCAAGCTTGTATGACCAAGCACTTCTGATGTTTGAAGTATCTGTTATC 420  
QY 421 TCTTCTGAAACCAATACAGACTAGATATTTGATGATGATGATGATGATGATGATG 480  
DB 421 TCTTCTGAAACCAATACAGACTAGATATTTGATGATGATGATGATGATGATGATG 480  
QY 481 AGCGGTTTCAATTTTATTTTATGAGAGCTTTTATGAGAGCTTTTATGAGAGCTTT 540  
DB 481 AGCGGTTTCAATTTTATTTTATGAGAGCTTTTATGAGAGCTTTTATGAGAGCTTT 540  
QY 541 ATAGGTTTATGATGCGGTATACAACTTAAACGATACCACTTTTATGAGAGCTTT 600  
DB 541 ATAGGTTTATGATGCGGTATACAACTTAAACGATACCACTTTTATGAGAGCTTT 600  
QY 601 GCGGCACTCTTCCGCTACAGACCTTTTACCAATTTTCTGCTGCTGCTGCTGCTGCT 660  
DB 601 GCGGCACTCTTCCGCTACAGACCTTTTACCAATTTTCTGCTGCTGCTGCTGCTGCT 660  
QY 661 GTACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
DB 661 GTACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
QY 721 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
DB 721 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
QY 781 TTCAACACGCTGCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 781 TTCAACACGCTGCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
DB 841 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 TCTGCTGCTTCAATTTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
DB 901 TCTGCTGCTTCAATTTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
QY 961 GGTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 961 GGTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

QY 1021 GCTCTGCGCAATATGACAGAGACAGAGTGGGCGGCTTGCCTAGCCGACCGCAGAGC 1080  
DB 1021 GCTCTGCGCAATATGACAGAGACAGAGTGGGCGGCTTGCCTAGCCGACCGCAGAGC 1080  
QY 1081 GGGGCTTGTTCGCGCGCTGAGCTTGCCTGCGGAGAGGAGACACCCGCTAGGGGGCCAGCTC 1140  
DB 1081 GGGGCTTGTTCGCGCGCTGAGCTTGCCTGCGGAGAGGAGACACCCGCTAGGGGGCCAGCTC 1140  
QY 1141 CAGCGCCGAG 1183  
DB 1141 CAGCGCCGAG 1183

RESULT 2  
US-10-790-430-8

; Sequence 8, Application US/10790430  
; Publication No. US2004013943A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMT32(ink603) and Composition and Methods for Detection of Invention  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/10/790,430  
; PRIOR FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 8  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(1183)  
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
; OTHER INFORMATION: 165-381 construct vector DNA  
; OTHER INFORMATION: 382-586 Zea maize plastid genes, rps11 and rpsa  
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA  
US-10-790-430-8

Query Match 100.0%; Score 1183; DB 18; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATAGCG 60  
DB 1 GACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATAGCG 60  
QY 61 GATGAAAAAATAATATAGCGCGCAAACTAGATATTAATTCGCGCGGCTCATCTAT 120  
DB 61 GATGAAAAAATAATATAGCGCGCAAACTAGATATTAATTCGCGCGGCTCATCTAT 120  
QY 121 GTTACTAGATCGGGGATATCCCGGGGAATTCGATACCAAGCTTTATATAGTAAAA 180  
DB 121 GTTACTAGATCGGGGATATCCCGGGGAATTCGATACCAAGCTTTATATAGTAAAA 180  
QY 181 GAGTAAATTTACATTTGGGCACTTTTATACGATATTTTACTTTATACACCTTTTA 240  
DB 181 GAGTAAATTTACATTTGGGCACTTTTATACGATATTTTACTTTATACACCTTTTA 240  
QY 241 ACTGATGTTTACATTTTGACCAAGTATCTTACTTTGTTTATTTTGACATATCCGA 300

```
Db 241 ACTGATGTTTACCTTTGACCGAGTATCTTACCTTTGTTATTTGACATCCCA 300
Qy 301 CTCTCTTCCAGCATATGATGACCTCGAGTAAAGCTTTGTAACGGCGCCCTAGGGA 360
Db 301 CTCTCTTCCAGCATATGATGATACCTCGAGTAAAGCTTTGTAACGGCGCCCTAGGGA 360
Qy 361 TATCAAGCTTGTACCAACCGGACACATTCACCTTATAGTGTGATGATCTCTTATC 420
Db 361 TATCAAGCTTGTACCAACCGGACACATTCACCTTATAGTGTGATGATCTCTTATC 420
Qy 421 TCTTCTCGAACCATTAACAGCTAGTATATTTGATGATGATGATGATGATGATGAT 480
Db 421 TCTTCTCGAACCATTAACAGCTAGTATATTTGATGATGATGATGATGATGATGAT 480
Qy 481 AGCGGTTTCAATTTTATTTTACAGAGCTTTTATTTTATTTTATTTTATTTTATTT 540
Db 481 AGCGGTTTCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 540
Qy 541 ATAGGTGTTTATCATCGCTATATACAACTTAACTGATACACCTTTTATGATGATGAT 600
Db 541 ATAGGTGTTTATCATCGCTATATACAACTTAACTGATACACCTTTTATGATGATGAT 600
Qy 601 GCGGATCTCTTCCGCTACCAAGCCTTTTACCACTTCTGCTGCTGCTGCTGCTGCT 660
Db 601 GCGGATCTCTTCCGCTACCAAGCCTTTTACCACTTCTGCTGCTGCTGCTGCTGCT 660
Qy 661 GTACGATATGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GTACGATATGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 ATAAATGATGACCAACCAACCTTGCCTTCACTACCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 ATAAATGATGACCAACCAACCTTGCCTTCACTACCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 TTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 TTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 841 GAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 GAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 901 TCTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 TCTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 961 GGTGCGGCGGCGCGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1020
Db 961 GGTGCGGCGGCGCGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1020
Qy 1021 GCTCTGCGCAATATCAAGAAACAGTGGGCGGCTGCGCTGAGCCGACCGCAGAGC 1080
Db 1021 GCTCTGCGCAATATCAAGAAACAGTGGGCGGCTGCGCTGAGCCGACCGCAGAGC 1080
Qy 1081 GGGGCTTGTGCGAGCCGTAGGCTGCGAGAGGGAAGGGAAGGGAAGGGAAGGGAAG 1140
Db 1081 GGGGCTTGTGCGAGCCGTAGGCTGCGAGAGGGAAGGGAAGGGAAGGGAAGGGAAG 1140
Qy 1141 CAGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
Db 1141 CAGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1183
```

```
RESULT 3
US-10-425-115-58081/c
; Sequence 58081, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 58081
; LENGTH: 6683
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRL4577_152966C.1
US-10-425-115-58081
Query Match 25.8%; Score 305; DB 18; Length 6683;
Best Local Similarity 100.0%; Pred. No. 3,1e-149; Indels 0; Gaps 0;
Matches 305; Conservative 0; Mismatches 0;
Qy 382 ACACACTTCACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
Db 5673 ACACACTTCACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5614
Qy 442 TAGATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
Db 5613 TAGATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5554
Qy 502 AGACGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 561
Db 5553 AGACGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5494
Qy 562 CAACCTTAACGATACACCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 621
Db 5493 CAACCTTAACGATACACCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5434
Qy 622 GCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 681
Db 5433 GCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5374
Qy 682 GTTCT 686
Db 5373 GTTCT 5369
```

```
RESULT 4
US-10-213-791-27/c
; Sequence 27, Application US/10213791
; Publication No. US20030106096A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15503)
; CURRENT APPLICATION NUMBER: US/10/213,791
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US/09/441,340
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/108,763
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression
; OTHER INFORMATION: cassette comprising a plant promoter linked to an
; OTHER INFORMATION: inton, a sequence encoding an AMPA acetyl
; OTHER INFORMATION: transferase, and a termination sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (28)..(965)
; FEATURE:
; NAME/KEY: inton
```

LOCATION: (966) ..(1423)  
FEATURE:  
NAME/KEY: transit\_peptide  
LOCATION: (1440) ..(1667)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1668) ..(2099)  
FEATURE:  
NAME/KEY: terminator  
LOCATION: (2114) ..(2369)  
US-10-213-791-27

Query Match 17.1%; Score 202; DB 15; Length 2378;  
Best Local Similarity 100.0%; Pred. No. 4.2e-95;  
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCGACCTTTATTAACGATATTT 222  
DB 202 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCGACCTTTATTAACGATATTT 143  
QY 223 ACTTATACACCTTTTAAGTATGTTTCACTTTTGACGAGTATCTTACCTTTGTT 282  
DB 142 ACTTATACACCTTTTAAGTATGTTTCACTTTTGACGAGTATCTTACCTTTGTT 83  
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTA 342  
DB 82 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTA 23  
QY 343 ACGCGCGCGCCCTTAGGATATC 364  
DB 22 ACGCGCGCGCCCTTAGGATATC 1

RESULT 5  
US-10-012-070A-49/c  
Sequence 49; Application US/10012070A  
Publication No. US20030077801A1  
GENERAL INFORMATION:  
APPLICANT: Hawkes, Timothy  
APPLICANT: Warner, Simon  
APPLICANT: Andrews, Christopher  
APPLICANT: Bachoo, Satvinder  
APPLICANT: Pickerill, Andrew  
TITLE OF INVENTION: Herbicide Resistant Plants  
FILE REFERENCE: 50490/UST  
CURRENT APPLICATION NUMBER: US/10/012,070A  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: PCT/GB00/01573  
PRIOR FILING DATE: 2000-04-20  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn Ver. 2.0  
SEO ID NO 49  
LENGTH: 1501  
TYPE: DNA  
ORGANISM: Oryza sp.  
US-10-012-070A-49

Query Match 14.5%; Score 171; DB 14; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 8.2e-79;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCGACCTTTATTAACGATATTT 222  
DB 920 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCGACCTTTATTAACGATATTT 861  
QY 223 ACTTATACACCTTTTAAGTATGTTTCACTTTTGACGAGTATCTTACCTTTGTT 282  
DB 860 ACTTATACACCTTTTAAGTATGTTTCACTTTTGACGAGTATCTTACCTTTGTT 801  
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTA 333  
DB 800 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTA 750

RESULT 6  
US-10-012-013-44/c  
Sequence 44; Application US/10012013  
Publication No. US20030079246A1  
GENERAL INFORMATION:  
APPLICANT: Hawkes, Timothy  
APPLICANT: Warner, Simon  
APPLICANT: Andrews, Christopher  
APPLICANT: Bachoo, Satvinder  
APPLICANT: Pickerill, Andrew  
TITLE OF INVENTION: Herbicide Resistant Plants  
FILE REFERENCE: 50450/UST  
CURRENT APPLICATION NUMBER: US/10/012,013  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: PCT/GB00/01572  
PRIOR FILING DATE: 2000-04-20  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.0  
SEO ID NO 44  
LENGTH: 1501  
TYPE: DNA  
ORGANISM: Oryza sp.  
US-10-012-013-44

Query Match 14.5%; Score 171; DB 14; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 8.2e-79;  
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCGACCTTTATTAACGATATTT 222  
DB 920 TTTTATATAGTAGAAGAGTAATTTCACTTTGGGCGACCTTTATTAACGATATTT 861  
QY 223 ACTTATACACCTTTTAAGTATGTTTCACTTTTGACGAGTATCTTACCTTTGTT 282  
DB 860 ACTTATACACCTTTTAAGTATGTTTCACTTTTGACGAGTATCTTACCTTTGTT 801  
QY 283 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTA 333  
DB 800 TATTTGGACTATCCGACTCTCTTCTCAAGCATATGATGACCTCGAGTATGTTA 750

RESULT 7  
US-10-678-588A-1  
Sequence 1; Application US/10678588A  
Publication No. US2005002266A1  
GENERAL INFORMATION:  
APPLICANT: Wu, Jingrui  
TITLE OF INVENTION: Water-Deficit-TolerantTransgenic Plants  
FILE REFERENCE: 38-21(52578)C  
CURRENT APPLICATION NUMBER: US/10/678,588A  
CURRENT FILING DATE: 2003-10-02  
PRIOR APPLICATION NUMBER: US 60/415,758  
PRIOR FILING DATE: 2002-10-02  
PRIOR APPLICATION NUMBER: US 60/425,157  
PRIOR FILING DATE: 2002-11-08  
PRIOR APPLICATION NUMBER: US 60/463,787  
PRIOR FILING DATE: 2003-04-11  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn version 3.2  
SEO ID NO 1  
LENGTH: 2480  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: transcriptional unit comprising promoter, coding sequence for  
OTHER INFORMATION: transcription factor of SEQ ID NO:2 and terminator elements  
US-10-678-588A-1

Query Match 14.3%; Score 169; DB 19; Length 2480;  
Best Local Similarity 100.0%; Pred. No. 9.2e-78;  
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

Db      6383 TTTTAAATAGTAGAGAAAAGAGTAATTTCACTTTGGGCCACCTTTTATTTACCATATTTT 6442
Oy      223 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGACAGGATATCTTACCTTTGTT 282
Db      6443 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGACAGGATATCTTACCTTTGTT 6502
Oy      283 TATTTGGACTATCCGACCTCTCTTCTCAAGCATATGATGAC 325
Db      6503 TATTTGGACTATCCGACCTCTCTCTCAAGCATATGATGAC 6545

RESULT 10
US-09-845-064-21
; Sequence 21, Application US/09845064
; Publication No. US20030175976A1
; GENERAL INFORMATION:
; APPLICANT: MERISTEM THERAPEUTICS
; TITLE OF INVENTION: CLEAN SYNTHETIC VECTORS, PLASMIDS, TRANSGENIC PLANTS
; TITLE OF INVENTION: AND PLANT PARTS CONTAINING SAID VECTORS, AND THEIR
; FILE REFERENCE: SYNVECI
; CURRENT APPLICATION NUMBER: US/09/845,064
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 10003
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
; FEATURE:
; NAME/KEY: rep_origin
; LOCATION: (1)-..(654)
; OTHER INFORMATION: Origin of replication ori RK2
; FEATURE:
; NAME/KEY: rep_origin
; LOCATION: (655)..(1263)
; OTHER INFORMATION: Origin of replication ori COLEI
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1264)..(2603)
; OTHER INFORMATION: NPT III gene coding for neomycin
; OTHER INFORMATION: phosphotransferase and kanamycin resistance
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2604)..(4098)
; OTHER INFORMATION: trfA locus from RK2 coding for two proteins, P285
; OTHER INFORMATION: and P382, enabling the increase of the replication
; OTHER INFORMATION: rate
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4106)..(4271)
; OTHER INFORMATION: T-DNA left border
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (4272)..(4559)
; OTHER INFORMATION: MCS multiple cloning site
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4594)..(5169)
; OTHER INFORMATION: Bar gene coding for phosphinothricine
; OTHER INFORMATION: acetyltransferase and glutosinate resistance
; FEATURE:
; NAME/KEY: intron
; LOCATION: (5170)..(5704)
; OTHER INFORMATION: Rice Actin Intron
; FEATURE:

```

```

; NAME/KEY: Promoter
; LOCATION: (5705)..(6638)
; OTHER INFORMATION: Rice Actin Promoter
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6646)..(6672)
; OTHER INFORMATION: MCS multiple cloning site
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (6672)..(7107)
; OTHER INFORMATION: High Molecular Weight Glutenin promoter from wheat
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7107)..(7169)
; OTHER INFORMATION: MCS multiple cloning site
; FEATURE:
; NAME/KEY: intron
; LOCATION: (7169)..(7687)
; OTHER INFORMATION: Rice Actin Intron
; FEATURE:
; NAME/KEY: gene
; LOCATION: (7688)..(9496)
; OTHER INFORMATION: GUS gene coding for beta glucuronidase
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (9497)..(9823)
; OTHER INFORMATION: No. US20030175976A1:aline synthetase terminator
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9823)..(9996)
; OTHER INFORMATION: T-DNA right border
; US-09-845-064-21

```

```

Query Match      13.8%; Score 163; DB 10; Length 10003;
Best Local Similarity 100.0%; Pred. No. 1,3e-74;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      163 TTTTAAATAGTAGAGAAAAGAGTAATTTCACTTTGGGCCACCTTTTATTTACCATATTTT 222
Db      6402 TTTTAAATAGTAGAGAAAAGAGTAATTTCACTTTGGGCCACCTTTTATTTACCATATTTT 6461
Oy      223 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGACAGGATATCTTACCTTTGTT 282
Db      6462 ACTTATATACCACTTTTAACTGATGTTTCACTTTTGACAGGATATCTTACCTTTGTT 6521
Oy      283 TATTTGGACTATCCGACCTCTCTTCTCAAGCATATGATGAC 325
Db      6522 TATTTGGACTATCCGACCTCTCTTCTCAAGCATATGATGAC 6564

RESULT 11
US-10-480-223A-23
; Sequence 23, Application US/10480223A
; Publication No. US20040250317A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto, Technology LLC
; APPLICANT: Huber, Scott A.
; APPLICANT: Doherty, Sean C.
; APPLICANT: Roberts, James K.
; APPLICANT: Shapley, Zachary W.
; TITLE OF INVENTION: Cotton Event 15985 and Compositions and Methods for Detection
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 11899,0232.PCUS00
; CURRENT APPLICATION NUMBER: US/10/480,223A
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US 60/297406
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PCT/US 02/17853
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 349

```

```

; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-480-223A-23

Query Match      12.4%; Score 147; DB 18; Length 349;
Best Local Similarity 100.0%; Pred. No. 3.3e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGTGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
    |||
Db 179 GACGTTATTTATGAGTGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 238

Qy 61 GATGAAAAACAATAATATAGCGCGCAAACTAGAGTAATATATCGCGCGGTGCATCTAT 120
    |||
Db 239 GATGAAAAACAATAATATAGCGCGCAAACTAGAGTAATATATCGCGCGGTGCATCTAT 298

Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
    |||
Db 299 GTTACTAGATCGGGGATATCCCGGGG 325

RESULT 12
US-10-480-223A-12
; Sequence 12, Application US/10480223A
; Publication No. US20040250317A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto, Technology LLC
; APPLICANT: Huber, Scott A
; APPLICANT: Doherty, Sean C.
; APPLICANT: Roberts, James K.
; APPLICANT: Shapley, Zachary W.
; TITLE OF INVENTION: Cotton Event 15985 and Compositions and Methods for Detection
; FILE REFERENCE: 11899.0232.PCUS00
; CURRENT APPLICATION NUMBER: US/10/480.223A.
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US 60/297406
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PCT/US 02/17853
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(349)
; OTHER INFORMATION: 3' end insert sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (350)..(673)
; OTHER INFORMATION: 3' end Gossypium hirsutum remnant DNA sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674)..(1360)
; OTHER INFORMATION: 3' end Gossypium hirsutum chromosomal sequence
US-10-480-223A-12

Query Match      12.4%; Score 147; DB 18; Length 1360;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGTGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
    |||
Db 179 GACGTTATTTATGAGTGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 238

Qy 61 GATGAAAAACAATAATATAGCGCGCAAACTAGAGTAATATATCGCGCGGTGCATCTAT 120
    |||
Db 239 GATGAAAAACAATAATATAGCGCGCAAACTAGAGTAATATATCGCGCGGTGCATCTAT 298

Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
```

```

Db 299 GTTACTAGATCGGGGATATCCCGGGG 325

RESULT 13
US-10-232-665-15
; Sequence 15, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232.665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377.466
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 3754
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (25)..(640)
; OTHER INFORMATION: P-CamV.35S
; FEATURE:
; NAME/KEY: intron
; LOCATION: (669)..(1472)
; OTHER INFORMATION: I-Zm.Hsp70
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1490)..(3448)
; OTHER INFORMATION: Cry3Bb1 variant v11231
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (3475)..(3730)
; OTHER INFORMATION: Agrobacterium tumefaciens nos 3' transcription
; OTHER INFORMATION: termination and polyadenylation sequence
US-10-232-665-15

Query Match      12.4%; Score 147; DB 15; Length 3754;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGTGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 60
    |||
Db 3601 GACGTTATTTATGAGTGGGTTTATGATTAGATCCCGCAATTATACATTAAATACGC 3660

Qy 61 GATGAAAAACAATAATATAGCGCGCAAACTAGAGTAATATATCGCGCGGTGCATCTAT 120
    |||
Db 3661 GATGAAAAACAATAATATAGCGCGCAAACTAGAGTAATATATCGCGCGGTGCATCTAT 3720

Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147
    |||
Db 3721 GTTACTAGATCGGGGATATCCCGGGG 3747

RESULT 14
US-10-232-665-13
; Sequence 13, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232.665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377.466
```



```
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 13
/ LENGTH: 4149
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: expression
/ FEATURE:
/ OTHER INFORMATION: cassette
/ FEATURE:
/ NAME/KEY: Promoter
/ LOCATION: (25)..(640)
/ OTHER INFORMATION: P-CanV.35S
/ FEATURE:
/ NAME/KEY: Intron
/ LOCATION: (669)..(1472)
/ OTHER INFORMATION: I-Zm.Hsp70
/ FEATURE:
/ NAME/KEY: transit_peptide
/ LOCATION: (1489)..(1635)
/ OTHER INFORMATION: amino terminal TS-Zm.rbcS
/ FEATURE:
/ NAME/KEY: Intron
/ LOCATION: (1636)..(1798)
/ OTHER INFORMATION: I-Zm.rbcS
/ FEATURE:
/ NAME/KEY: transit_peptide
/ LOCATION: (1799)..(1885)
/ OTHER INFORMATION: carboxy terminus TS-Zm.rbcS
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1885)..(3843)
/ OTHER INFORMATION: Cry3Bb1 variant v11231
/ FEATURE:
/ NAME/KEY: terminator
/ LOCATION: (3871)..(4127)
/ OTHER INFORMATION: T-AGRTu.nos 3' transcription termination and
/ OTHER INFORMATION: polyadenylation sequence
/ US-10-232-665-13
```

```
Query Match 12.4%; Score 147; DB 15; Length 4149;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 60
DB 3996 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 4055
QY 61 GATGAAAACAAATATAGCGCGCAACTAGATTAATTTATCGCGCGGTGTCATCTAT 120
DB 4056 GATGAAAACAAATATAGCGCGCAACTAGATTAATTTATCGCGCGGTGTCATCTAT 4115
QY 121 GTTACTGATCGGGGATATCCCGGGG 147
DB 4116 GTTACTGATCGGGGATATCCCGGGG 4142
```

## RESULT 15

```
US-10-198-478-16
/ Sequence 16, Application US/10198478
/ Publication No. US20030188336A1
/ GENERAL INFORMATION:
/ APPLICANT: Corbin, David R.
/ APPLICANT: Romano, Charles P.
/ TITLE OF INVENTION: Improved Methods for Transforming Plants to Express delta-Endotox
/ FILE REFERENCE: 38-21 (15547) B
/ CURRENT APPLICATION NUMBER: US/10/198,478
/ PRIOR FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: 09/186, 002
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 16
```

```
/ LENGTH: 8349
/ TYPE: DNA
/ ORGANISM: artificial sequence
/ FEATURE:
/ OTHER INFORMATION: completely synthesized
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(8349)
/ OTHER INFORMATION:
/ US-10-198-478-16
```

```
Query Match 12.4%; Score 147; DB 16; Length 8349;
Best Local Similarity 100.0%; Pred. No. 3.4e-66;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 60
DB 5722 GACGTTATTTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC 5781
QY 61 GATGAAAACAAATATAGCGCGCAACTAGATTAATTTATCGCGCGGTGTCATCTAT 120
DB 5782 GATGAAAACAAATATAGCGCGCAACTAGATTAATTTATCGCGCGGTGTCATCTAT 5841
QY 121 GTTACTGATCGGGGATATCCCGGGG 147
DB 5842 GTTACTGATCGGGGATATCCCGGGG 5868
```

```
Search completed: February 9, 2005, 11:12:28
Job time : 1969.4 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 18.6055 Seconds  
(without alignments)  
5727.097 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18  
Sequence: 1 Tgcgtctctgcctgactt 18

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseq11980s:\*  
2: geneseq11990s:\*  
3: geneseq12000s:\*  
4: geneseq12001as:\*  
5: geneseq12001bs:\*  
6: geneseq12002as:\*  
7: geneseq12002bs:\*  
8: geneseq12003as:\*  
9: geneseq12003bs:\*  
10: geneseq12003cs:\*  
11: geneseq12003ds:\*  
12: geneseq12004as:\*  
13: geneseq12004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 18    | 100.0       | 18     | 6     | ABK15248            |
| 2          | 18    | 100.0       | 18     | 6     | ABK15248            |
| 3          | 16    | 88.9        | 1361   | 5     | AAK86560 DNA encod  |
| 4          | 16    | 88.9        | 1587   | 4     | AAH14053 Human CDN  |
| 5          | 15    | 83.3        | 21     | 3     | AAZ73249 Human bra  |
| 6          | 15    | 83.3        | 445    | 5     | ABV19889 Human pro  |
| 7          | 15    | 83.3        | 500    | 5     | ABV49652 Human pro  |
| 8          | 15    | 83.3        | 715    | 4     | ABL09089 Drosophila |
| 9          | 15    | 83.3        | 2280   | 10    | ADBS9493 Human gen  |
| 10         | 15    | 83.3        | 2280   | 10    | ADBS9493 Human gen  |
| 11         | 15    | 83.3        | 3167   | 4     | AAH06260 DNA encod  |
| 12         | 15    | 83.3        | 3167   | 4     | AAH06260 DNA encod  |
| 13         | 15    | 83.3        | 3190   | 4     | ABL09088 Drosophila |
| 14         | 15    | 83.3        | 5919   | 10    | ADBS9752 Human con  |
| 15         | 15    | 83.3        | 41335  | 4     | AAK86502 Human imm  |
| 16         | 15    | 83.3        | 62130  | 12    | ADQ97885 Mouse can  |
| 17         | 15    | 83.3        | 70019  | 13    | ABD33601 Murine ca  |
| 18         | 14    | 77.8        | 20     | 2     | AAK96131 PCR prime  |
| 19         | 14    | 77.8        | 25     | 12    | ADP16460 Renal cel  |
| 20         | 14    | 77.8        | 25     | 12    | ADP16461 Renal cel  |

|    |    |      |     |    |           |
|----|----|------|-----|----|-----------|
| 21 | 14 | 77.8 | 27  | 10 | ADD40762  |
| 22 | 14 | 77.8 | 149 | 5  | ABV15475  |
| 23 | 14 | 77.8 | 154 | 5  | ABV15397  |
| 24 | 14 | 77.8 | 169 | 4  | ABA47800  |
| 25 | 14 | 77.8 | 169 | 4  | AAK14094  |
| 26 | 14 | 77.8 | 169 | 6  | ABSI13933 |
| 27 | 14 | 77.8 | 230 | 5  | ABV36189  |
| 28 | 14 | 77.8 | 241 | 10 | ACA55805  |
| 29 | 14 | 77.8 | 241 | 12 | ADIS5601  |
| 30 | 14 | 77.8 | 300 | 2  | AAZ14788  |
| 31 | 14 | 77.8 | 366 | 8  | ACA44008  |
| 32 | 14 | 77.8 | 379 | 5  | ABV36267  |
| 33 | 14 | 77.8 | 400 | 9  | ACH46396  |
| 34 | 14 | 77.8 | 426 | 8  | ACA35231  |
| 35 | 14 | 77.8 | 459 | 11 | ACH94209  |
| 36 | 14 | 77.8 | 483 | 12 | ACH90762  |
| 37 | 14 | 77.8 | 496 | 4  | ABA42671  |
| 38 | 14 | 77.8 | 496 | 4  | AAK01345  |
| 39 | 14 | 77.8 | 496 | 5  | ABV45241  |
| 40 | 14 | 77.8 | 496 | 5  | ABV45307  |
| 41 | 14 | 77.8 | 496 | 6  | ABE01398  |
| 42 | 14 | 77.8 | 515 | 6  | ABN79319  |
| 43 | 14 | 77.8 | 543 | 4  | AAH70725  |
| 44 | 14 | 77.8 | 547 | 12 | ACH77045  |
| 45 | 14 | 77.8 | 549 | 3  | AAK93637  |

#### ALIGNMENTS

RESULT 1  
ID ABK15248 standard; DNA: 18 BP.  
AC ABK15248;  
DT 08-MAY-2002 (first entry)  
XX  
DE Corn transgene junction probe #3.  
XX  
KW Probe; ss; rice actin 1 promoter; RAI; RAI intron;  
KW chloroplast transit peptide gene; glyphosate resistance; corn;  
KW 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
KW transcripional terminator; cauliflower mosaic virus 35S promoter;  
KW PV-ZMGF32; transgenic; nk603.  
XX  
OS Zea mays.  
OS Synthetic.  
XX  
PN EPI167531-A1.  
XX  
PD 02-JAN-2002.  
XX  
PF 15-JUN-2001; 2001EP-00202314.  
XX  
PR 22-JUN-2000; 2000US-0213567P.  
PR 13-OCT-2000; 2000US-0240014P.  
PR 13-OCT-2000; 2000US-0241215P.  
XX  
PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
FI Behr CF, Hitonaka C, Heck GR, You J;  
XX  
DR WPI; 2002-165871/22.  
XX  
PT Novel DNA construct useful for producing a corn plant that tolerates  
PT application of glyphosate herbicide, comprises two transgene expres-  
PT cassettes.  
XX  
PS Claim 8; Page 3; 25pp; English.  
XX  
CC The invention relates to a DNA construct (I) comprising a first (F1) and  
CC a second (F2) expression cassette. F1 of the DNA construct, in operable

CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
CC PV-ZMG12(nk603) of 498 or 1183bp, or primers and probes derived from  
CC them (used to detect transgene junction points). The construct is useful  
CC for producing a corn plant that tolerates application of glyphosate  
CC herbicide by transforming a corn cell with, selecting the corn cell into a  
CC tolerance to application of glyphosate, and growing the corn cell into a  
CC fertile corn plant. The present sequence is a transgene junction probe  
CC sequence consisting of corn genomic DNA and vector sequences  
CC  
SQ Sequence 18 BP; 1 A; 4 C; 4 G; 9 T; 0 U; 0 Other;  
Query Match 100.0%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Gaps 0;  
Matches 18; Conservative 0; Indels 0; Indels 0; Gaps 0;  
Oy 1 TGCTGTTCTGCTGACTTT 18  
1 TGCTGTTCTGCTGACTTT 18  
Db 1 TGCTGTTCTGCTGACTTT 18  
RESULT 2  
ABK15245  
ID ABK15245 standard; DNA; 1183 BP.  
XX  
AC ABK15245;  
XX  
XX 08-MAY-2002 (first entry)  
DE Corn nk603 3' transgene/genomic sequence PCR fragment.  
XX  
XX  
KM ds; rice actin 1 promoter; RA1; RA1 intron;  
KM chloroplast transit peptide gene; glyphosate resistance; corn;  
KM 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
KM transcriptional terminator; cauliflower mosaic virus 35S promoter;  
KM PV-ZMG12; transgenic; nk603.  
XX  
XX  
OS Zea mays.  
OS Escherichia coli.  
OS Agrobacterium tumefaciens.  
OS Synthetic.  
OS Chimeric.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1..164  
FT misc\_feature /\*tag= a  
FT misc\_feature /note= "T-AGRTU.nos vector sequence"  
FT misc\_feature 165..381  
FT misc\_feature /\*tag= b  
FT misc\_feature /note= "Vector sequence"  
FT misc\_feature 382..686  
FT misc\_feature /\*tag= c  
FT misc\_feature /note= "Corn plasmid genes rps11/rpoA partial sequence"  
FT misc\_feature 687..1183  
FT misc\_feature /\*tag= d  
FT misc\_feature /note= "Corn genomic sequence"  
XX  
XX  
XX EPI167531-A1.  
XX  
XX  
XX 02-JAN-2002.  
XX  
XX  
XX 15-JUN-2001; 2001EB-00202314.  
XX  
XX  
XX 22-JUN-2000; 2000US-0213567P.  
XX 13-OCT-2000; 2000US-0240014P.  
XX 13-OCT-2000; 2000US-0241215P.  
XX  
XX (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
XX Behr CF, Hironaka C, Heck GR, You J;  
PI

XX  
DR WPI; 2002-165871/22.  
XX  
XX Novel DNA construct useful for producing a corn plant that tolerates  
PT application of glyphosate herbicide, comprises two transgene expression  
PT cassettes.  
XX  
XX  
PS Claim 3; Page 16-17; 25pp; English.  
XX  
XX The invention relates to a DNA construct (I) comprising a first (F1) and  
CC a second (F2) expression cassette. F1 of the DNA construct, in operable  
CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
CC PV-ZMG12(nk603) of 498 or 1183bp, or primers and probes derived from corn  
CC them (used to detect transgene junction points). The construct is useful  
CC for producing a corn plant that tolerates application of glyphosate  
CC herbicide by transforming a corn cell with, selecting the corn cell for  
CC tolerance to application of glyphosate, and growing the corn cell into a  
CC fertile corn plant. The present sequence is the 1183bp amplicon  
CC consisting of the Agrobacterium transcriptional terminator sequence,  
CC vector sequences and corn sequences  
CC  
SQ Sequence 1183 BP; 295 A; 289 C; 319 T; 0 U; 0 Other;  
Query Match 100.0%; Score 18; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 2.1; Mismatches 0; Gaps 0;  
Matches 18; Conservative 0; Indels 0; Indels 0; Gaps 0;  
Oy 1 TGCTGTTCTGCTGACTTT 18  
1 TGCTGTTCTGCTGACTTT 695  
Db 1 TGCTGTTCTGCTGACTTT 695  
RESULT 3  
AAS66560  
ID AAS66560 standard; cDNA; 1361 BP.  
XX  
AC AAS66560;  
XX  
XX 13-FEB-2002 (first entry)  
DE  
DE DNA encoding novel human diagnostic protein #2364.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KM Food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
PD  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX P-PSDB; ABG02373.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 2364; 103bp; English.  
PS

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed CC genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC useful for generating antibodies against it, detecting or quantitating a CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic CC coding sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1361 BP; 414 A; 322 C; 313 G; 313 T; 0 U; 0 Other;

XX Query Match 88.9%; Score 16; DB 5; Length 1361;

XX Best Local Similarity 100.0%; Pred. No. 27;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Db 1 TGCTGTTCTGCTGACT 16

XX 948 TGCTGTTCTGCTGACT 963

#### RESULT 4

XX AAH14053 standard; cDNA; 1587 BP.

XX AC AAH14053;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:11181.

XX KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX PS Claim 8; SEQ ID NO 11181; 2537bp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: CC (a) an oligo-dt primer and an oligonucleotide complementary to the CC complementary strand of a polynucleotide which comprises one of the 5602 CC nucleotide sequences defined in the specification; where the CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination CC of an oligonucleotide comprising a sequence complementary to the CC complementary strand of a polynucleotide which comprises a 5'-end CC sequence and an oligonucleotide comprising a sequence complementary to a CC polynucleotide which comprises a 3'-end sequence; where the CC oligonucleotide comprises at least 15 nucleotides and the combination of CC the 5'-end sequence/3'-end sequence is selected from those defined in the CC specification. The primer sets can be used in antisense therapy and in CC gene therapy. The primers are useful for synthesizing polynucleotides, CC particularly full-length cDNAs. The primers are also useful for the CC detection and/or diagnosis of the abnormality of the proteins encoded by CC the full-length cDNAs. The primers allow obtaining of the full-length CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and CC AAH1633 to AAH18742 represent human cDNA sequences; AA892446 to AA895893 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent CC oligonucleotides, all of which are used in the exemplification of the CC present invention

XX Sequence 1587 BP; 312 A; 416 C; 383 G; 476 T; 0 U; 0 Other;

XX Query Match 88.9%; Score 16; DB 4; Length 1587;

XX Best Local Similarity 100.0%; Pred. No. 27;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Db 1 TGCTGTTCTGCTGACT 16

XX 948 TGCTGTTCTGCTGACT 963

#### RESULT 5

XX AAZ73249 standard; DNA; 21 BP.

XX AC AAZ73249;

XX DT 10-SEP-2001 (first entry)

XX DE Human biallelic marker upstream amplification primer SEQ ID NO:7605.

XX KM Human genome; biallelic marker; high density disequilibrium map;

XX KM genomic map; haplotype; phenotype; polymorphic base; genotyping;

XX KM haplotyping; hybridisation; identification; characterisation;

XX KM amplification; single nucleotide polymorphism; SNP; PCR primer;

XX KM diagnosis; ss.

XX OS Homo sapiens.

XX PN WO954500-A2.

XX PD 28-OCT-1999.

XX PF 21-APR-1999; 99WO-IB000822.

XX PF 21-APR-1998; 98US-0082614P.

XX PR 23-NOV-1998; 98US-0109732P.

XX PA (GEST ) GENSET.

XX PI Cohen D, Blumenfeld M, Chumakov I;

XX DR WPI; 2000-013267/01.

XX PT Novel biallelic markers used to construct a high density disequilibrium map of the human genome.

XX PS Claim 9; Page 1852; 2745pp; English.



Query Match 83.3%; Score 15; DB 5; Length 500;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 TGCTGTTCTGCTGAC 15  
Db 296 TGCTGTTCTGCTGAC 282

## RESULT 8

ABL09089

ID ABL09089 standard; cDNA; 715 BP.

AC ABL09089;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 21749.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW,

XX WPI; 2001-656860/75.

XX P-PSDB; ABB64986.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.

XX Claim 1; SEQ ID NO 21749; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 715 BP; 151 A; 192 C; 202 G; 170 T; 0 U; 0 Other;

XX Query Match 83.3%; Score 15; DB 4; Length 715;  
XX Best Local Similarity 100.0%; Pred. No. 91;  
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 TGCTGTTCTGCTGAC 15

XX 88 TGCTGTTCTGCTGAC 102

## RESULT 9

ADE59493/C

ID ADE59493 standard; DNA; 2280 BP.

XX ADE59493;

XX 29-JAN-2004 (first entry)  
XX Human gene AK001865, SEQ ID NO 5388.  
XX Human; de; gene; pain; neuronal tissue; gene therapy;  
XX spinal segmental nerve injury; chronic constriction injury; CCI;  
XX spared nerve injury; SN1; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

XX (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AK001865.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
XX or human polynucleotides or a polynucleotide which represents a fragment,  
XX derivative or allelic variation of the nucleic acid sequence. Also  
XX claimed are a vector comprising the novel polynucleotide, a host cell  
XX comprising the vector, a method for identifying a nucleotide sequence  
XX which is differentially regulated in an animal subjected to pain and a  
XX kit to perform the method, an array, a method for identifying an agent  
XX that increases or decreases the expression of the polynucleotide sequence  
XX that is differentially expressed in neuronal tissue of a first animal  
XX subjected to pain, a method for identifying a compound which regulates  
XX the expression of a polynucleotide sequence which is differentially  
XX expressed in an animal subjected to pain, a method for identifying a  
XX compound that regulates the activity of one or more of the  
XX polynucleotides, a method for producing a pharmaceutical composition, a  
XX method for identifying a method for producing a pharmaceutical composition, a  
XX method for identifying a method for producing a pharmaceutical composition, a  
XX method for identifying a method for producing a pharmaceutical composition,  
XX activity in an animal of one or more of the polypeptides given in the  
XX specification, a method for identifying a compound useful in treating  
XX pain and a pharmaceutical composition comprising the one or more  
XX polypeptides or their antibodies. The polynucleotide or the compound that  
XX modulates its activity is useful for preparing a medicament for treating  
XX pain (e.g. spinal segmental nerve injury (SN1), chronic constriction  
XX injury (CCI) and spared nerve injury (SN1) in an animal (e.g. gene  
XX therapy). The sequence presented is a human DNA (shown in Table 2 of the  
XX specification) which encodes one of the polypeptides of the invention  
XX which is differentially expressed during pain. Note: The sequence data  
XX for this patent did not form part of the printed specification, but was  
XX obtained in electronic form directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2280 BP; 680 A; 358 C; 392 G; 850 T; 0 U; 0 Other;

XX Query Match 83.3%; Score 15; DB 10; Length 2280;  
XX Best Local Similarity 100.0%; Pred. No. 97;  
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 4 TGTTCTGCTGACTTT 18

XX 741 TGTTCTGCTGACTTT 727

RESULT 10  
ADES9491/c  
ID ADE59491 standard; DNA; 2280 BP.  
XX  
AC ADE59491;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human gene AK001865, SEQ ID NO 5386.  
XX  
KW Human; ds; gene; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003;  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GENO ) GEN HOSPITAL CORP.  
XX (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M,  
XX  
DR WPI; 2003-268312/26.  
XX GENBANK; AK001865.  
PT  
PT New composition comprising two or more isolated polypeptides, useful for  
XX preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017p; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
XX or human polynucleotides or a polynucleotide which represents a fragment,  
XX derivative or allelic variation of the nucleic acid sequence. Also  
XX claimed are a vector comprising the novel polynucleotide, a host cell  
XX comprising the vector, a method for identifying a nucleotide sequence  
XX which is differentially regulated in an animal subjected to pain and a  
XX kit to perform the method, an array, a method for identifying an agent  
XX that increases or decreases the expression of the polynucleotide sequence  
XX that is differentially expressed in neuronal tissue of a first animal  
XX subjected to pain, a method for identifying a compound which regulates  
XX the expression of a polynucleotide sequence which is differentially  
XX expressed in an animal subjected to pain, a method for identifying a  
XX compound that regulates the activity of one or more of the  
XX polynucleotides, a method for producing a pharmaceutical composition, a  
XX method for identifying a compound or small molecule that regulates the  
XX activity in an animal of one or more of the polypeptides given in the  
XX specification, a method for identifying a compound useful in treating  
XX pain and a pharmaceutical composition comprising the one or more  
XX polypeptides or their antibodies. The polynucleotide or the compound that  
XX modulates its activity is useful for preparing a medicament for treating  
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
XX therapy). The sequence presented is a human DNA (shown in Table 2 of the  
XX specification) which encodes one of the polypeptides of the invention  
XX which is differentially expressed during pain. Note: The sequence data  
XX for this patent did not form part of the printed specification, but was  
XX obtained in electronic form directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 2280 BP; 680 A; 358 C; 392 G; 850 T; 0 U; 0 Other;  
XX  
XX Query Match 83.3%; Score 15; DB 10; Length 2280;  
XX Best Local Similarity 100.0%; Pred. No. 97;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 4 TGTTCGTGACTT 18  
Db 741 TGTTCGTGACTT 727  
RESULT 11  
AAS30480  
ID AAS30480 standard; DNA; 3167 BP.  
XX  
AC AAS30480;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE DNA encoding novel prostate gland antigen, Seq ID NO 338.  
XX  
XX Human; nocotropic; neuroprotective; cytostatic; antiapoptotic;  
KW antianemic; dermatological; immunosuppressive; antiinflammatory;  
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;  
KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;  
KW prostatic; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;  
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;  
KW reproductive system disorder; autoimmune disorder; urinary system;  
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;  
KW blood-related disorder; hyperproliferative disorder; respiratory;  
KW neurological disorder; endocrine disorder; inflammatory disorder;  
KW liver disorder; wound healing; food preservative; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200155447-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001330.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-019076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 14-AUG-2000; 2000US-0225799P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.



RESULT 12  
AAL06260  
ID AAL06260 standard; DNA; 3167 BP.  
XX  
AC AAL06260;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human reproductive system related antigen DNA SEQ ID NO: 8948.  
XX  
KW Human; reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001339.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236337P.  
PR 29-SEP-2000; 2000US-0236337P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249256P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.



PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P..  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251688P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-46570/50.  
DR  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
XX  
XX Disclosure; SEQ ID NO 8948; 1297bp + Sequence Listing; English.  
PS  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention  
XX  
SQ Sequence 3167 BP; 892 A; 608 C; 749 G; 918 T; 0 U; 0 Other;  
XX  
Query Match 83.3%; Score 15; DB 4; Length 3167;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 TGTTCTGCTGACTTT 18  
Db 1739 TGTTCTGCTGACTTT 1753  
XX  
RESULT 13  
ABL09088/c  
ID ABL09088 standard; cDNA; 3190 BP.  
XX  
XX ABL09088;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 21746.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical; gene; ss.  
XX  
XX Drosophila melanogaster.  
OS  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US009221.  
PF  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
XX  
XX P-PSDB; ABB64985.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT Interactions.  
XX  
XX Claim 1; SEQ ID NO 21746; 21bp + Sequence Listing; English.  
PS  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-AB130511), expressed DNA  
CC sequences (ABL161840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX  
SQ Sequence 3190 BP; 967 A; 700 C; 674 G; 849 T; 0 U; 0 Other;  
XX  
Query Match 83.3%; Score 15; DB 4; Length 3190;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCTGTTCTGCTGAC 15  
Db 2024 TGCTGTTCTGCTGAC 2010  
XX  
RESULT 14  
ADP59752/c  
ID ADP59752 standard; cDNA; 5919 BP.  
XX  
XX ADP59752;  
AC  
XX  
XX 12-FEB-2004 (first entry)  
DT  
XX  
XX Human contig polynucleotide sequence SEQ ID NO:2119.  
DE  
XX  
XX biological activity; genetic engineering; hybridisation probe; oligomer;  
KM primer; chromosome mapping; gene mapping; recombinant protein production;  
KW human; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2003080795-A2.  
PN  
XX  
XX 02-OCT-2003.  
PD  
XX  
XX 09-AUG-2002; 2002WO-US025485.  
PF  
XX  
XX 09-AUG-2001; 2001US-0311261P.  
PR  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Yang Y, Wang Z, Wang G, Ma Y;  
PI  
XX WPI; 2003-876918/81.  
DR  
XX  
XX P-PSDB; ADF60204.  
DR  
XX  
XX New polynucleotides, useful as hybridization probes, oligomers or  
PT primers, for chromosome or gene mapping, for the recombinant production  
PT of proteins, and for generating antisense DNA or RNA.  
XX  
XX Example 2; SEQ ID NO 2119; 571bp; English.  
PS  
XX  
XX The present invention describes isolated polynucleotide sequences (I),  
CC which encode polypeptides (II) with biological activity. Also described:  
CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)  
CC a host cell genetically engineered to comprise (I) which is operatively  
CC associated with a regulatory sequence that modulates expression of (I) in  
CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition  
CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed  
CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of  
CC (4) in a sample; (8) identifying a compound that binds to the polypeptide  
CC of (4); (9) producing the polypeptide of (4); and (10) a collection of

CC polynucleotides comprising at least one of the polynucleotide sequences  
CC (1). The polynucleotides (1) can be used as hybridisation probes,  
CC oligomers or primers, for chromosome or gene mapping, for the recombinant  
CC production of proteins, and for generating antisense DNA or RNA. The  
CC present sequence represents a human contig polynucleotide sequence, which  
CC is used in an example from the present invention.

Sequence 5919 BP; 1846 A; 1292 C; 1202 G; 1579 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 10; Length 5919;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TGTTCGTGACTT 18

Db 5871 TGTTCGTGACTT 5857

RESULT 15

AAK6202  
ID AAK6202 standard; DNA; 41335 BP.

AC AAK6202;

DT 07-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41014.

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

cytostatic; gene therapy; vaccine; metastasis; ds.

Homo sapiens.

PN WO200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US001354.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 29-SEP-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239355P.

PR 13-OCT-2000; 2000US-0239357P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.



**This Page Blank (uspto)**

121 GTTACTAGATCGGGATATCCCGGGAAATCGGATCAAGCTTTTAAATAGTAAAA 180  
181 GAGTAAATTTCACTTTGGGCACTTTTATACCATTTTATACCACTTTT 240  
181 GAGTAAATTTCACTTTGGGCACTTTTATACCATTTTATACCACTTTT 240  
241 ACTGATGTTTTCACCTTTGACCAAGTAACTTAACTTTGTTTATTTGACTATCCGA 300  
241 ACTGATGTTTTCACCTTTGACCAAGTAACTTAACTTTGTTTATTTGACTATCCGA 300  
301 CTCTCTTTCAGACATATGATTAATCTTCAGATGCTTGTAAAGGCGGCGCTTAAAGGA 360  
301 CTCTCTTTCAGACATATGATTAATCTTCAGATGCTTGTAAAGGCGGCGCTTAAAGGA 360  
361 TATCAAGCTTGTACCAAGCACTTCCACTAGTGTGAGTGAATCCTGTATC 420  
361 TATCAAGCTTGTACCAAGCACTTCCACTAGTGTGAGTGAATCCTGTATC 420  
421 TCTTCTGCAACATTAACAAGATTAATTTATGATCAATGAACTGTTATTTCTTGA 480  
421 TCTTCTGCAACATTAACAAGATTAATTTATGATCAATGAACTGTTATTTCTTGA 480  
481 AGCGGTTTCACTTTTATTAACAAGCTTTTATGAGAGTGAATCCTATATCGGC 540  
481 AGCGGTTTCACTTTTATTAACAAGCTTTTATGAGAGTGAATCCTATATCGGC 540  
541 ATAGGTGTTACATGCGCTATTAACAATTAACCTTATGAGATGCTGCTAT 600  
541 ATAGGTGTTACATGCGCTATTAACAATTAACCTTATGAGATGCTGCTAT 600  
601 GCGGCATCTCTTCCGCTACCAAGCACTTTTACATTAATCTTCTGCTGCAAAACCCT 660  
601 GCGGCATCTCTTCCGCTACCAAGCACTTTTACATTAATCTTCTGCTGCAAAACCCT 660  
661 GTAGGAATAGCATATGATGCTGCTTCTGCTGCTGCTTATTTTAAATGTAAGGA 720  
661 GTAGGAATAGCATATGATGCTGCTTCTGCTGCTGCTTATTTTAAATGTAAGGA 720  
721 ATAAATATGACAAACAACCTTCCCTTCACTACCGGTGCGAGGCAAGAGTGGG 780  
721 ATAAATATGACAAACAACCTTCCCTTCACTACCGGTGCGAGGCAAGAGTGGG 780  
781 TTCAACACGCTGCGACACGATGCAACGACCTTCCAGCAATCTCGAGGCGGAC 840  
781 TTCAACACGCTGCGACACGATGCAACGACCTTCCAGCAATCTCGAGGCGGAC 840  
841 GAGCACTGAGCAGGCTGCGATTAACAAGCTGCGGCACTTCTTCTTCTT 900  
841 GAGCACTGAGCAGGCTGCGATTAACAAGCTGCGGCACTTCTTCTTCTT 900  
901 TCTCTGCTTCAACTTGGCGCGGAGTGTCTAGACCCAGGAGATGCTGTGAGAGA 960  
901 TCTCTGCTTCAACTTGGCGCGGAGTGTCTAGACCCAGGAGATGCTGTGAGAGA 960  
961 GGTGCGGCGGCGCAATTTTATAGCTGAGGCAAGCAAGCTTGGCGCAATCCAGATCCAGA 1020  
961 GGTGCGGCGGCGCAATTTTATAGCTGAGGCAAGCAAGCTTGGCGCAATCCAGATCCAGA 1020  
1021 GCTCTGCGCAATATCAAGAAACAAGTGGGCGCTGCGCTTACCCAGCCGCAAGAGC 1080  
1021 GCTCTGCGCAATATCAAGAAACAAGTGGGCGCTGCGCTTACCCAGCCGCAAGAGC 1080  
1081 GGGGCTTGTGCGAGCGCTAGGCTGCGGAGGAGGAGCCGCTAGGGGGGCGCATGCTC 1140  
1081 GGGGCTTGTGCGAGCGCTAGGCTGCGGAGGAGGAGCCGCTAGGGGGGCGCATGCTC 1140  
1141 CAGCGCCAGAGAGAAAAAGAAAGAAAGGCGGAGATGATG 1183  
1141 CAGCGCCAGAGAGAAAAAGAAAGAAAGGCGGAGATGATG 1183

RESULT 2  
US-09-441-340-27/c

Sequence 27, Application US/09441340  
Patent No. 6448476  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
TITLE OF INVENTION: Phosphonate Metabolizing Plants  
FILE REFERENCE: 38-21(15303)  
CURRENT FILING DATE: US/09/441,340  
EARLIER APPLICATION NUMBER: 60/108,763  
EARLIER FILING DATE: 1998-11-17  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 27  
LENGTH: 2378  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: expression  
OTHER INFORMATION: cassette comprising a plant promoter linked to an  
OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl  
OTHER INFORMATION: transferase, and a termination sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (28) .. (965)  
FEATURE:  
NAME/KEY: intron  
LOCATION: (966) .. (1423)  
FEATURE:  
NAME/KEY: transist peptide  
LOCATION: (1440) .. (1667)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1668) .. (2099)  
FEATURE:  
NAME/KEY: terminator  
LOCATION: (2114) .. (2369)  
US-09-441-340-27

Query Match 17.1%; Score 202; DB 3; Length 2378;  
Best Local Similarity 100.0%; Pred. No. 2.5e-89;  
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTTTATATATGTAAGAAAGATTAATTTCACTTTGGGCACTTTTATACGATATTTT 222  
DB 202 TTTTATATATGTAAGAAAGATTAATTTCACTTTGGGCACTTTTATACGATATTTT 143  
QY 223 ACTTATACCACTTTTATCTGATGTTTCACTTTGACAGGATATCTTACCTTGT 282  
DB 142 ACTTATACCACTTTTATCTGATGTTTCACTTTGACAGGATATCTTACCTTGT 83  
QY 283 TATTTGACTATCCGACTCTTCTCAAGCATATGAATGACTGAGTAACTTGT 342  
DB 82 TATTTGACTATCCGACTCTTCTCAAGCATATGAATGACTGAGTAACTTGT 23  
QY 343 ACGGCGCGGCGCTAGGATATC 364  
DB 22 ACGGCGCGGCGCTAGGATATC 1

RESULT 3  
US-09-068-101-5/c  
Sequence 5, Application US/09068101  
Patent No. 642960  
GENERAL INFORMATION:  
APPLICANT: PLANT GENETIC SYSTEMS N.V.  
TITLE OF INVENTION: Improved Barstar Gene  
FILE REFERENCE: 2121-1139P  
CURRENT FILING DATE: US/09/068,101  
EARLIER APPLICATION NUMBER: 60/108,763  
EARLIER FILING DATE: 1998-08-26  
EARLIER APPLICATION NUMBER: EP 96202446.9  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent Ver. 2.0

This Page Blank (uspto)

Qy 61 GATAGAAAACAAATATATGCGCGCAACTAGATTAATTTATCGCGCGGTGCATCTAT 120  
Db 3022 GATAGAAAACAAATATATGCGCGCAACTAGATTAATTTATCGCGCGGTGCATCTAT 3081  
Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147  
Db 3082 GTTACTAGATCGGGGATATCCCGGGG 3108

RESULT 9  
US-09-186-002-15  
Sequence 15, Application US/09186002B  
Patent No. 6489542  
GENERAL INFORMATION:  
APPLICANT: Romano, Charles P.  
APPLICANT: Corbin, David R.  
TITLE OF INVENTION: Improved Method for Transforming Plants to Express  
FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002  
CURRENT FILING DATE: 1998-11-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 15  
LENGTH: 10252  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (3660)..(3773)  
OTHER INFORMATION: "n" = g, a, c, or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4355)..(4407)  
OTHER INFORMATION: "n" = g, a, c, or t  
US-09-186-002-15

Query Match 12.4%; Score 147; DB 4; Length 10252;  
Best Local Similarity 100.0%; Pred. No. 2.9e-62;  
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAAGATCCCGCAATTATACATTAAATACGC 60  
Db 2965 GACGTTATTTATGAGATGGGTTTATGATTAAGATCCCGCAATTATACATTAAATACGC 3024  
Qy 61 GATAGAAAACAAATATATGCGCGCAACTAGATTAATTTATCGCGCGGTGCATCTAT 120  
Db 3025 GATAGAAAACAAATATATGCGCGCAACTAGATTAATTTATCGCGCGGTGCATCTAT 3084  
Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147  
Db 3085 GTTACTAGATCGGGGATATCCCGGGG 3111

RESULT 10  
US-09-186-002-13  
Sequence 13, Application US/09186002B  
Patent No. 6489542  
GENERAL INFORMATION:  
APPLICANT: Romano, Charles P.  
APPLICANT: Corbin, David R.  
TITLE OF INVENTION: Improved Method for Transforming Plants to Express  
FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002  
CURRENT FILING DATE: 1998-11-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 13  
LENGTH: 10339  
TYPE: DNA  
ORGANISM: Artificial Sequence

FEATURE:  
NAME/KEY: unsure  
LOCATION: (3687)..(3760)  
OTHER INFORMATION: "n" = g, a, c, or t  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4382)..(4434)  
OTHER INFORMATION: "n" = g, a, c, or t  
US-09-186-002-13

Query Match 12.4%; Score 147; DB 4; Length 10339;  
Best Local Similarity 100.0%; Pred. No. 2.9e-62;  
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAAGATCCCGCAATTATACATTAAATACGC 60  
Db 3052 GACGTTATTTATGAGATGGGTTTATGATTAAGATCCCGCAATTATACATTAAATACGC 3111  
Qy 61 GATAGAAAACAAATATATGCGCGCAACTAGATTAATTTATCGCGCGGTGCATCTAT 120  
Db 3112 GATAGAAAACAAATATATGCGCGCAACTAGATTAATTTATCGCGCGGTGCATCTAT 3171  
Qy 121 GTTACTAGATCGGGGATATCCCGGGG 147  
Db 3172 GTTACTAGATCGGGGATATCCCGGGG 3198

RESULT 11  
US-09-441-340-27  
Sequence 27, Application US/09441340  
Patent No. 6489542  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
TITLE OF INVENTION: Phosphonate Metabolizing Plants  
FILE REFERENCE: 38-21(15303)  
CURRENT FILING DATE: 1998-11-16  
EARLIER APPLICATION NUMBER: 60/108,763  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 27  
LENGTH: 2378  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: expression  
OTHER INFORMATION: cassette comprising a plant promoter linked to an  
OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl  
OTHER INFORMATION: transferase, and a termination sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (28)..(965)  
FEATURE:  
NAME/KEY: intron  
LOCATION: (966)..(1423)  
FEATURE:  
NAME/KEY: transic peptide  
LOCATION: (1440)..(1667)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1668)..(2099)  
FEATURE:  
NAME/KEY: terminator  
LOCATION: (2114)..(2369)  
US-09-441-340-27

Query Match 11.8%; Score 140; DB 3; Length 2378;  
Best Local Similarity 100.0%; Pred. No. 8.5e-59;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGTTATTTATGAGATGGGTTTATGATTAAGATCCCGCAATTATACATTAAATACGC 60

**This Page Blank (uspto)**



Mon Feb 14 10:13:01 2005

us-10-790-430-8.01igo.rml

Page 6

Db 2239 GACGTTATTATGAGATGGGTTTATGATAGTCCGCAATTATTAATTAATACCC 2298  
QY 61 GATAGAAAACAAATATAGCGCCCAACTAGATTAATTTATCGCGCGGTGCATCTAT 120  
Db 2299 GATAGAAAACAAATATAGCGCCCAACTAGATTAATTTATCGCGCGGTGCATCTAT 2358  
QY 121 GTTACTAGATCGGGGATATC 140  
Db 2359 GTTACTAGATCGGGGATATC 2378

RESULT 12  
US-09-441-340-31  
Sequence 31, Application US/09441340  
Patent No. 6448476  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
TITLE OF INVENTION: Phosphonate Metabolizing Plants  
FILE REFERENCE: 38-21(15303)  
CURRENT FILING DATE: 1999-11-16  
EARLIER APPLICATION NUMBER: 60/108,763  
EARLIER FILING DATE: 1998-11-17  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 31  
LENGTH: 2436  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: monocot  
OTHER INFORMATION: expression cassette comprising plant operable  
OTHER INFORMATION: promoter linked to an intron, a sequence coding  
OTHER INFORMATION: for an AMPA acetyltransferase, and a termination  
OTHER INFORMATION: sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (26)..(640)  
FEATURE:  
NAME/KEY: intron  
LOCATION: (670)..(1473)  
FEATURE:  
NAME/KEY: transit peptide  
LOCATION: (1498)..(1725)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1726)..(2157)  
FEATURE:  
NAME/KEY: terminator  
LOCATION: (2172)..(2427)  
US-09-441-340-31

Query Match 11.8%; Score 140; DB 3; Length 2436;  
Best Local Similarity 100.0%; Pred. No. 8.5e-59;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGTTATTATGAGATGGGTTTATGATAGTCCGCAATTATTAATTAATACCC 60  
Db 2297 GACGTTATTATGAGATGGGTTTATGATAGTCCGCAATTATTAATTAATACCC 2356  
QY 61 GATAGAAAACAAATATAGCGCCCAACTAGATTAATTTATCGCGCGGTGCATCTAT 120  
Db 2357 GATAGAAAACAAATATAGCGCCCAACTAGATTAATTTATCGCGCGGTGCATCTAT 2416  
QY 121 GTTACTAGATCGGGGATATC 140  
Db 2417 GTTACTAGATCGGGGATATC 2436

RESULT 13  
US-09-441-340-23  
Sequence 23, Application US/09441340  
Patent No. 6448476

GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
TITLE OF INVENTION: Phosphonate Metabolizing Plants  
FILE REFERENCE: 38-21(15303)  
CURRENT FILING DATE: 1999-11-16  
EARLIER APPLICATION NUMBER: 60/108,763  
EARLIER FILING DATE: 1998-11-17  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 23  
LENGTH: 1630  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: expression  
OTHER INFORMATION: cassette comprising plant operable promoter linked  
OTHER INFORMATION: to a coding sequence encoding an AMPA  
OTHER INFORMATION: acetyltransferase linked to a transcription  
OTHER INFORMATION: termination sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (33)..(605)  
FEATURE:  
NAME/KEY: transit peptide  
LOCATION: (627)..(892)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (893)..(1324)  
FEATURE:  
NAME/KEY: terminator  
LOCATION: (1350)..(1605)  
US-09-441-340-23

Query Match 11.6%; Score 137; DB 3; Length 1630;  
Best Local Similarity 100.0%; Pred. No. 2.6e-57;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACGTTATTATGAGATGGGTTTATGATAGTCCGCAATTATTAATTAATACCC 60  
Db 1476 GACGTTATTATGAGATGGGTTTATGATAGTCCGCAATTATTAATTAATACCC 1535  
QY 61 GATAGAAAACAAATATAGCGCCCAACTAGATTAATTTATCGCGCGGTGCATCTAT 120  
Db 1536 GATAGAAAACAAATATAGCGCCCAACTAGATTAATTTATCGCGCGGTGCATCTAT 1595  
QY 121 GTTACTAGATCGGGGAT 137  
Db 1596 GTTACTAGATCGGGGAT 1612

RESULT 14  
US-09-098-219B-5  
Sequence 5, Application US/09098219B  
Patent No. 6441277  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard  
APPLICANT: Cheikh, No. 6441277dine  
APPLICANT: Kishore, Ganesh  
TITLE OF INVENTION: Expression of Fructose 1,6 Biphosphate  
TITLE OF INVENTION: Aldolase in Transgenic Plants  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Arnold White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: US  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC compatible  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 121.68 Seconds  
(Without alignments)  
5630.828 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 1 taccacgcgcacacttc 18

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hlc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_g881.\*  
9: gb\_g882.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 88.9  | 1441        | 9      | AG076194 | AG076194 Pan trogl |
| 2          | 83.3  | 247         | 9      | CL445530 | CL445530 ZMMBB046  |
| 3          | 83.3  | 364         | 5      | BU067975 | BU067975 1632_B05  |
| 4          | 83.3  | 370         | 5      | BU068295 | BU068295 2556_F04  |
| 5          | 83.3  | 399         | 5      | BU068642 | BU068642 2562_H04  |
| 6          | 83.3  | 414         | 5      | BU068903 | BU068903 2569_B09  |
| 7          | 83.3  | 440         | 6      | CB039236 | CB039236 AG_U3_31G |
| 8          | 83.3  | 449         | 8      | A2141658 | A2141658 SP_0003_B |
| 9          | 83.3  | 452         | 1      | AL924785 | AL924785 AL924785  |
| 10         | 83.3  | 453         | 5      | BU068182 | BU068182_2555_B04  |
| 11         | 83.3  | 456         | 7      | CK989655 | CK989655 BgHC-2_38 |
| 12         | 83.3  | 459         | 6      | CA954936 | CA954936 K132e04.Y |
| 13         | 83.3  | 477         | 4      | BI593866 | BI593866 AG_U3_06A |
| 14         | 83.3  | 481         | 5      | BU067587 | BU067587 1626_F12  |
| 15         | 83.3  | 483         | 5      | CA953980 | CA953980 K137a10.Y |
| 16         | 83.3  | 498         | 4      | BM515610 | BM515610 K151d06.Y |
| 17         | 83.3  | 503         | 4      | BM516183 | BM516183 K174e03.Y |
| 18         | 83.3  | 508         | 4      | BM515619 | BM515619 K151e03.Y |
| 19         | 83.3  | 509         | 5      | BU066631 | BU066631 1609_B02  |
| 20         | 83.3  | 517         | 6      | CB039563 | CB039563 AG_U3_37H |
| 21         | 83.3  | 524         | 4      | BM283870 | BM283870 K128f01.Y |
| 22         | 83.3  | 524         | 4      | BM568782 | BM568782 K141h07.Y |
| 23         | 83.3  | 524         | 6      | CA849289 | CA849289 K110c10.Y |
| 24         | 83.3  | 524         | 6      | CA953971 | CA953971 K124h07.Y |

|      |    |      |     |   |          |                    |
|------|----|------|-----|---|----------|--------------------|
| C 25 | 15 | 83.3 | 524 | 6 | CA954253 | CA954253 K140e12.Y |
| C 26 | 15 | 83.3 | 524 | 6 | CA954961 | CA954961 K132h07.Y |
| C 27 | 15 | 83.3 | 524 | 6 | CA954983 | CA954983 K134c02.Y |
| C 28 | 15 | 83.3 | 527 | 2 | BE604361 | BE604361 WHE1405-1 |
| C 29 | 15 | 83.3 | 537 | 6 | CA954745 | CA954745 K130a03.Y |
| C 30 | 15 | 83.3 | 540 | 6 | CA953797 | CA953797 K122c05.Y |
| C 31 | 15 | 83.3 | 546 | 6 | BM283807 | BM283807 K152g03.Y |
| C 32 | 15 | 83.3 | 547 | 6 | CD883084 | CD883084 F1.112E01 |
| C 33 | 15 | 83.3 | 549 | 5 | BU067422 | BU067422 1620_G02  |
| C 34 | 15 | 83.3 | 549 | 6 | CA954309 | CA954309 K141c06.Y |
| C 35 | 15 | 83.3 | 560 | 6 | BM281999 | BM281999 K124b01.Y |
| C 36 | 15 | 83.3 | 560 | 6 | CA849866 | CA849866 K106d02.Y |
| C 37 | 15 | 83.3 | 562 | 4 | BM280788 | BM280788 K105a07.Y |
| C 38 | 15 | 83.3 | 562 | 6 | CA954933 | CA954933 K132d12.Y |
| C 39 | 15 | 83.3 | 567 | 4 | BM568585 | BM568585 K139b05.Y |
| C 40 | 15 | 83.3 | 569 | 6 | CA953755 | CA953755 K121e11.Y |
| C 41 | 15 | 83.3 | 570 | 4 | BM567304 | BM567304 K101b03.Y |
| C 42 | 15 | 83.3 | 571 | 6 | CA850315 | CA850315 K126g05.Y |
| C 43 | 15 | 83.3 | 572 | 4 | BM569276 | BM569276 K146f07.Y |
| C 44 | 15 | 83.3 | 572 | 6 | CA850345 | CA850345 K127b05.Y |
| C 45 | 15 | 83.3 | 572 | 6 | CA850345 | CA850345 K127b05.Y |

## ALIGNMENTS

|            |   |   |  |         |     |        |                 |
|------------|---|---|--|---------|-----|--------|-----------------|
| RESULT 1   | AG076194  | LOCUS   | AG076194   | 1441 bp | DNA | linear | GSS 03-NOV-2001 |
| DEFINITION | Pan troglodytes DNA, clone: PTB-070F05.R, genomic survey sequence.  | ACCESSION   | AG076194   |         |     |        |                 |
| VERSION    | AG076194.1  | GI  | 16627996   |         |     |        |                 |
| KEYWORDS   | GSS.  | SOURCE  | Pan troglodytes (chimpanzee)   |         |     |        |                 |
| ORGANISM   | Pan troglodytes   |   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan. |         |     |        |                 |
| REFERENCE  | 1   | Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. |  |         |     |        |                 |
| TITLE      | BAC end sequences of library PTB  |   |  |         |     |        |                 |
| JOURNAL    | Unpublished   |   |  |         |     |        |                 |
| AUTHORS    | 2 (bases 1 to 1441)   |   |  |         |     |        |                 |
| TITLE      | Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170) |   |  |         |     |        |                 |
| REFERENCE  | Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.   |   |  |         |     |        |                 |
| PRIMERS    | LIBRARY   |   |  |         |     |        |                 |
|            | Sequencing: M13rev  |   |  |         |     |        |                 |
|            | Vector : pKS145   |   |  |         |     |        |                 |
|            | R.Site 1 : SacI   |   |  |         |     |        |                 |
|            | R.Site 2 : SacI   |   |  |         |     |        |                 |
| FEATURES   | Location/Qualifiers   |   |  |         |     |        |                 |
| source     | 1..1441   |   |  |         |     |        |                 |
|            | /organism="Pan troglodytes"   |   |  |         |     |        |                 |
|            | /mol_type="genomic DNA"   |   |  |         |     |        |                 |
|            | /db_xref="taxon:9598"   |   |  |         |     |        |                 |
|            | /clone="PTB-070F05.R"   |   |  |         |     |        |                 |
|            | /sex="male"   |   |  |         |     |        |                 |
|            | /cell_type="lymphoblast"  |   |  |         |     |        |                 |
|            | /clone_lib="PTB Chimpanzee Male BAC Library"  |   |  |         |     |        |                 |
| ORIGIN     | Query Match   |   |  |         |     |        |                 |
|            | Best local similarity   |   |  |         |     |        |                 |
|            | 88.9%; Score 16; DB 9; Length 1441;   |   |  |         |     |        |                 |
|            | 100.0%; Pred. No. 14;   |   |  |         |     |        |                 |

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGGACACACT 16  
 |||||  
 Db 1136 TACCACGGACACACT 1151

RESULT 2  
 LOCUS CL445530/c 247 bp DNA linear GSS 24-MAR-2004  
 DEFINITION ZMWB0463D18f ZMWB0463D18f (HindIII) Zea mays genomic clone  
 ZMWB0463D18 5', genomic survey sequence.  
 ACCESION CL445530  
 VERSION CL445530  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 247)  
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,  
 Zhovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.  
 Sequencing of the maize genome at PGR (2003c)  
 Unpublished (2003)  
 Contact: Bharti,A.K.  
 Dr.Jochim Messing's lab  
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers  
 University  
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
 Tel: 732 445 3801  
 Fax: 732 445 5735  
 Email: bharti@waksman.rutgers.edu  
 Seq primer: T7  
 Class: BAC ends  
 High quality sequence start: 111.  
 Location/Qualifiers  
 1..247  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultiivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMWB0463D18"  
 /lab\_host="E. coli DH10B"  
 /clone\_id="ZMWB0463D18f (HindIII)"  
 /note="Vector: pCUG1; Site\_1: HindIII; Site\_2: HindIII"

ORIGIN

Query Match 83.3%; Score 15; DB 9; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 Db 49 CACGCGACACACTTC 35

RESULT 3  
 LOCUS BU067975 364 bp mRNA linear EST 27-AUG-2002  
 DEFINITION 1632 E05 J106T5 Mature perithecia *Gibberella zeae* cDNA, mRNA  
 sequence.  
 ACCESION BU067975  
 VERSION BU067975.1 GI:22509164  
 KEYWORDS EST.  
 SOURCE *Gibberella zeae*  
 ORGANISM *Gibberella zeae*  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocerales; Nectriaceae; *Gibberella*.  
 1 (bases 1 to 364)  
 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.  
 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.  
 Analysis of expressed sequence tags from *Gibberella zeae* (anamorph  
*Fusarium graminearum*)

JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)  
 MEDLINE 22508120  
 PUBMED 12620255  
 COMMENT Contact: Frances Trail  
 Department of Plant Biology  
 Michigan State University  
 East Lansing, MI 48824, USA  
 Tel: 517 432 2939  
 Fax: 517 353 1926  
 Email: trail@msu.edu.

FEATURES  
 source  
 1..364  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="NRRL 31084"  
 /db\_xref="taxon:5518"  
 /clone\_id="Mature perithecia"  
 /note="Vector: Ziplox; Site\_1: NotI; Site\_2: SalI"

ORIGIN

Query Match 83.3%; Score 15; DB 5; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 Db 299 CACGCGACACACTTC 313

RESULT 4  
 LOCUS BU068295 370 bp mRNA linear EST 27-AUG-2002  
 DEFINITION 2556 F04\_K08ZT5 Mature perithecia *Gibberella zeae* cDNA, mRNA  
 sequence.  
 ACCESION BU068295  
 VERSION BU068295.1 GI:22509484  
 KEYWORDS EST.  
 SOURCE *Gibberella zeae*  
 ORGANISM *Gibberella zeae*  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocerales; Nectriaceae; *Gibberella*.  
 1 (bases 1 to 370)  
 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.  
 Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C.  
 Analysis of expressed sequence tags from *Gibberella zeae* (anamorph  
*Fusarium graminearum*)  
 Fungal Genet. Biol. 38 (2), 187-197 (2003)

JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)  
 MEDLINE 22508120  
 PUBMED 12620255  
 COMMENT Contact: Frances Trail  
 Department of Plant Biology  
 Michigan State University  
 East Lansing, MI 48824, USA  
 Tel: 517 432 2939  
 Fax: 517 353 1926  
 Email: trail@msu.edu.

FEATURES  
 source  
 1..370  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="NRRL 31084"  
 /db\_xref="taxon:5518"  
 /clone\_id="Mature perithecia"  
 /note="Vector: Ziplox; Site\_1: NotI; Site\_2: SalI"

ORIGIN

RESULT 5  
 BU068642 399 bp mRNA linear EST 27-AUG-2002  
 LOCUS 2562\_H04\_P08275 Mature perithecia *Gibberella zeae* cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION BU068642  
 VERSION BU068642.1 GI:22509831  
 KEYWORDS EST.  
 SOURCE *Gibberella zeae*  
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; *Gibberella*.  
 REFERENCE 1 (bases 1 to 399)  
 AUTHORS Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.  
 TITLE Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*)  
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)  
 MEDLINE 22508120  
 PUBMED 12620255  
 COMMENT Contact: Frances Trail  
 Department of Plant Biology  
 Michigan State University  
 East Lansing, MI 48824, USA  
 Tel: 517 432 2939  
 Fax: 517 353 1926  
 Email: trail@msu.edu.  
 Location/Qualifiers  
 1..399  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="NRRL 31084"  
 /db\_xref="taxon:5518"  
 /clone\_idb="Mature perithecia"  
 /note="Vector: Ziplox; Site\_1: NotI; Site\_2: SalI"

ORIGIN  
 Query Match 83.3%; Score 15; DB 5; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 Db 300 CACGCGACACACTTC 314

RESULT 6  
 BU068903 414 bp mRNA linear EST 27-AUG-2002  
 LOCUS 2569\_E09\_J17ZT5 Mature perithecia *Gibberella zeae* cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION BU068903  
 VERSION BU068903.1 GI:22510092  
 KEYWORDS EST.  
 SOURCE *Gibberella zeae*  
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; *Gibberella*.  
 REFERENCE 1 (bases 1 to 414)  
 AUTHORS Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.  
 TITLE Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*)  
 JOURNAL Fungal Genet. Biol. 38 (2), 187-197 (2003)  
 MEDLINE 22508120  
 PUBMED 12620255  
 COMMENT Contact: Frances Trail  
 Department of Plant Biology  
 Michigan State University  
 East Lansing, MI 48824, USA  
 Tel: 517 432 2939  
 Fax: 517 353 1926  
 Email: trail@msu.edu.  
 Location/Qualifiers

FEATURES  
 source 1..414  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="NRRL 31084"  
 /db\_xref="taxon:5518"  
 /clone\_idb="Mature perithecia"  
 /note="Vector: Ziplox; Site\_1: NotI; Site\_2: SalI"

ORIGIN  
 Query Match 83.3%; Score 15; DB 6; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

source 1..414  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="NRRL 31084"  
 /db\_xref="taxon:5518"  
 /clone\_idb="Mature perithecia"  
 /note="Vector: Ziplox; Site\_1: NotI; Site\_2: SalI"

ORIGIN  
 Query Match 83.3%; Score 15; DB 5; Length 414;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 Db 299 CACGCGACACACTTC 313

RESULT 7  
 CB039236/c 440 bp mRNA linear EST 15-JAN-2003  
 LOCUS As L3\_31G08 SKPL *Ascaris suum* (parasitic nematode) L3 stage *Ascaris*  
 DEFINITION sum cDNA clone As L3\_31G08 5' similar to p41105 60S ribosomal  
 protein L28. Mus musculus (Mouse), mRNA sequence.  
 ACCESSION CB039236  
 VERSION CB039236.1 GI:27758480  
 KEYWORDS EST.  
 SOURCE *Ascaris suum* (pig roundworm)  
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascarididae; *Ascaris*.  
 REFERENCE 1 (bases 1 to 440)  
 AUTHORS Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Gulliano, D., Hall, N., Quayle, M. and Barrell, B.  
 TITLE Edinburgh University/Sanger Centre Nematode EST Project  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Blaxter ML  
 Institute of Cell, Animal and Population Biology  
 University of Edinburgh  
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK.  
 Tel: +44 131 650 6760  
 Fax: +44 131 670 5450  
 Email: mark.blaxter@ed.ac.uk  
 The library was prepared for Malcolm Kennedy by Joyce Moore,  
 Glasgow University. Sequencing was performed by the Pathogen  
 Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike  
 Quail & Bart Barrell).  
 PCR Primers  
 FORWARD: T3  
 BACKWARD: T7PL  
 Plate: 31 row: G column: 08  
 Seq primer: SKPL  
 High quality sequence stop: 296.  
 Location/Qualifiers  
 1..440  
 /organism="Ascaris suum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6253"  
 /clone="As L3\_31G08"  
 /sex="mixed"  
 /dev\_stage="L3"  
 /clone\_idb="Ascaris suum (parasitic nematode) L3 stage"  
 /note="Vector: Lambda ZAP II; Site\_1: EcoRI (5'end); The  
 Site\_2: (3'end); *Ascaris suum* is a parasite of pigs. The  
 library was constructed from mRNA from *Ascaris suum* L3  
 stage."

ORIGIN  
 Query Match 83.3%; Score 15; DB 6; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 405 CACGCGACACACTTC 391

Db

RESULT 8  
 A2141658 449 bp DNA linear GSS 28-AUG-2000  
 LOCUS SP 0003 B1 A08 SP66 Strongylocentrotus purpuratus, purple sea  
 DEFINITION urchin sperm genomic BAC library Strongylocentrotus purpuratus  
 genomic clone Plate=3 Col=15 Row=B, genomic survey sequence.

ACCESSION A2141658  
 VERSION A2141658.1 GI:8293561  
 KEYWORDS GSS.  
 SOURCE Strongylocentrotus purpuratus  
 ORGANISM Strongylocentrotus purpuratus  
 Eukaryote; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoida;  
 Strongylocentrotidae; Strongylocentrotus.

REFERENCE  
 1 (bases 1 to 449)  
 Cameron, R.A., Mahairas, G., Raab, J.P., Martinez, P., Biondi, T.R.,  
 Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T.,  
 Wray, G.A., Eitensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H.  
 and Hood, L.  
 A sea urchin genome project: Sequence scan, virtual map, and  
 additional resources  
 Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

TITLE  
 A sea urchin genome project: Sequence scan, virtual map, and  
 additional resources

JOURNAL  
 MEDLINE  
 PUBMED  
 20402566  
 10920195

COMMENT  
 Contact: Cameron, RA, Davidson, EH, Hood, L  
 Division of Biology 156-29  
 California Institute of Technology  
 Pasadena California 91125, USA  
 Tel: (626) 395-8421  
 Fax: (626) 793-3047  
 Email: acameron@caltech.edu  
 Plate: 3 row: B column: 15  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence atop: 449.

FEATURES  
 source  
 1..449  
 /organism="Strongylocentrotus purpuratus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7668"  
 /clone="Plate=3 Col=15 Row=B"  
 /clone\_lib="Strongylocentrotus purpuratus, purple sea  
 urchin, sperm genomic BAC library"  
 /note="Organ: Sperm; Vector: BACs.6; BAC Clones in E-Coli  
 DH10B"

ORIGIN

Query Match 83.3%; Score 15; DB 8; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 23 CACGCGACACACTTC 37

Db

RESULT 9  
 AL924785 452 bp mRNA linear EST 06-JUL-2004  
 LOCUS AL924785 PUR-Z1+Z2 Danio rerio cDNA clone 145-H10-2, mRNA sequence.  
 DEFINITION AL924785  
 ACCESSION AL924785  
 VERSION AL924785.1 GI:23191365  
 KEYWORDS EST.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.

REFERENCE  
 1 (bases 1 to 452)  
 Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Sun, A., He, Y., Ma, W.,  
 Wang, W., Wen, Z., and Peng, J.  
 15000 unique zebrafish EST clusters and their future use in  
 microarray for profiling gene expression patterns during  
 embryogenesis  
 Genome Res. 13 (3), 455-466 (2003)

JOURNAL  
 MEDLINE  
 PUBMED  
 22505427  
 12618376

COMMENT  
 Contact: Peng J  
 Lab of Functional Genomics  
 Institute of Molecular and Cell Biology  
 30 Medical Drive, Singapore, 117609, Singapore  
 Email: pengj@imcb.a-star.edu.sg  
 Clone requests: info@openbioystems.com  
 Open Bioystems,  
 6705 Odyssey Drive, Huntsville, AL 35806.

FEATURES  
 source  
 1..452  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /strain="local wildtype"  
 /db\_xref="taxon:7955"  
 /clone="145-H10-2"  
 /tissue\_type="whole embryo or fish"  
 /dev\_stage="mixed stages"  
 /clone\_lib="PUR-Z1+Z2"

ORIGIN

Query Match 83.3%; Score 15; DB 1; Length 452;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
 |||||  
 26 CACGCGACACACTTC 12

Db

RESULT 10  
 BU068182 453 bp mRNA linear EST 27-AUG-2002  
 LOCUS BU068182  
 DEFINITION 2555\_B04\_C07ZT5 Mature perithecia Gibberella zeae cDNA, mRNA  
 sequence.

ACCESSION BU068182  
 VERSION BU068182.1 GI:22509371  
 KEYWORDS EST.  
 SOURCE Gibberella zeae  
 ORGANISM Gibberella zeae  
 Eukaryote; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

REFERENCE  
 1 (bases 1 to 453)  
 Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.  
 Analysis of expressed sequence tags from Gibberella zeae (anamorph  
 Fusarium graminearum)  
 Fungal Genet. Biol. 38 (2), 187-197 (2003)

JOURNAL  
 MEDLINE  
 PUBMED  
 12620255  
 22508120

COMMENT  
 Contact: Frances Trail  
 Department of Plant Biology  
 Michigan State University  
 East Lansing, MI 48824, USA  
 Tel: 517 432 2939  
 Fax: 517 353 1926  
 Email: trail@msu.edu.

FEATURES  
 source  
 1..453  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="NRRL 31084"  
 /db\_xref="taxon:5518"  
 /clone\_lib="Mature perithecia"  
 /note="Vector: Ziplox; Site\_1: NotI; Site\_2: SalI"

ORIGIN

Query Match 83.3%; Score 15; DB 5; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACTTC 18  
 |||||  
 DB 299 CACGCGACACTTC 313

RESULT 11  
 LOCUS CK989655 456 bp mRNA linear EST 19-MAR-2004  
 DEFINITION BgHC-2.38 BgHC Biomphalaria glabrata cDNA 5', mRNA sequence.  
 ACCESSION CK989655  
 VERSION CK989655.1 GI:45596692  
 KEYWORDS EST.  
 SOURCE Biomphalaria glabrata (bloodfluke planorb)  
 ORGANISM Biomphalaria glabrata  
 Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata;  
 Basommatophora; Lymnaeidae; Planorbidae; Biomphalaria.  
 1 (bases 1 to 456)  
 Mitter, G., Gallinder, R., Tisseyre, P., Allienne, J.-F., Carmignani, C.  
 and Couesau, C.  
 ESTs from Biomphalaria glabrata hemocytes  
 Unpublished (2004)  
 Contact: Miltia Guillaume  
 Parasitologie fonctionnelle et evolutive.  
 CNRS / Univ. Perpignan  
 52 Ave Paul Alduy, Universite, 66 860 Perpignan cedex, France  
 Tel: 33 3 68 66 21 88  
 Fax: 33 3 68 66 22 81  
 Email: miltia@univ-perp.fr  
 Seq primer: CGAAGTATCTCGACG.  
 Location/Qualifiers  
 1. 456  
 /organism="Biomphalaria glabrata"  
 /mol\_type="mRNA"  
 /isolate="Bg. Bra strain"  
 /db\_xref="taxon:6526"  
 /sex="hermaphrodite"  
 /cell\_type="circulating hemocytes"  
 /clone\_id="BgHC"  
 /note="Vector: pDNR-LIB; The library was designed with 500  
 ng of total RNA using the creator SMART cDNA library  
 construction kit (Clontech)."

ORIGIN  
 Query Match 83.3%; Score 15; DB 7; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCGACACTT 17  
 |||||  
 DB 184 CCACGCGACACTT 198

RESULT 12  
 LOCUS CA954936 459 bp mRNA linear EST 02-JAN-2003  
 DEFINITION K132e04.y1 Ascaris suum embryo SL1 TOPO v1 Ascaris suum cDNA 5'  
 similar to SW:RL28\_MOUSE P41105 60S RIBOSOMAL PROTEIN L28. ;, mRNA  
 sequence.  
 ACCESSION CA954936  
 VERSION CA954936.1 GI:27471296  
 KEYWORDS EST.  
 SOURCE Ascaris suum (pig roundworm)  
 ORGANISM Ascaris suum  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae;  
 Ascaridoidea; Ascarididae; Ascaris.  
 1 (bases 1 to 459)  
 McCarter, J., Clifton, S., Chiappelli, B., Page, D., Martin, J.,  
 Wylie, T., Dance, M., Maitra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,  
 Tsagaris, J., R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,  
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,  
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.  
 The Washington Univ. Nematode EST Project, 1999  
 Unpublished (1999)  
 Contact: McCarter JP  
 The Washington Univ. Nematode EST Project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 The library was constructed by Claire Murphy and Dr. James McCarter  
 at Washington University, St. Louis. Oligo(dT)-SL1 PCR based  
 library. Embryo cDNA PCR products of size >400 nucleotides  
 containing SL1 on the 5' end and oligo(dT) on the 3' end were  
 non-directionally ligated into PCR11-TOPO(Invitrogen) following the  
 TOPO TA cloning protocol. 30-60 cell embryo material was provided  
 by Dr. Richard Davis of City University of New York Graduate  
 Center, College of Staten Island, Staten Island, NY  
 (redavis@postbox.csi.cuny.edu).  
 Seq primer: SL1 primer  
 High quality sequence stop: 437.  
 Location/Qualifiers  
 1. 459  
 /organism="Ascaris suum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6253"  
 /dev\_stage="30-60 cell embryo"  
 /lab\_host="DH10B"  
 /clone\_id="Ascaris suum embryo SL1 TOPO v1"  
 /note="Vector: PCR11-TOPO (Invitrogen); Site 1: EcoRI;  
 Site 2: EcoRI; The library was constructed by Claire  
 Murphy and Dr. James McCarter at Washington University,  
 St. Louis. Oligo(dT)-SL1 PCR based library. Embryo cDNA  
 PCR products of size >400 nucleotides containing SL1 on  
 the 5' end and oligo(dT) on the 3' end were  
 non-directionally ligated into PCR11-TOPO(Invitrogen)  
 following the TOPO TA cloning protocol. 30-60 cell embryo  
 material was provided by Dr. Richard Davis of City  
 University of New York Graduate Center, College of Staten  
 Island, Staten Island, NY (redavis@postbox.csi.cuny.edu)."

ORIGIN  
 Query Match 83.3%; Score 15; DB 6; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACTTC 18  
 |||||  
 DB 425 CACGCGACACTTC 411

RESULT 13  
 LOCUS B1593866 477 bp mRNA linear EST 07-SEP-2001  
 DEFINITION Ae L3 06A09 SKPL Ascaris suum (parasitic nematode) L3 stage Ascaris  
 suum cDNA clone Ae L3 06A09 5' similar to ref|NP\_033107.1|  
 ribosomal protein L28 - Mus musculus, mRNA sequence.  
 ACCESSION B1593866  
 VERSION B1593866.1 GI:15497353  
 KEYWORDS EST.  
 SOURCE Ascaris suum (pig roundworm)  
 ORGANISM Ascaris suum  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae;  
 Ascaridoidea; Ascarididae; Ascaris.  
 1 (bases 1 to 477)  
 Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guilianno, D.,  
 Hall, N., Quayle, M. and Barrell, B.  
 Edinburgh University/Sanger Centre Nematode EST Project  
 Unpublished (2000)

COMMENT  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared for Malcolm Kennedy by Joyce Moore,  
Glasgow University. Sequencing was performed by the Pathogen  
Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike  
Quail & Bart Barrell).  
PCR Primers  
FORWARD: T3  
BACKWARD: T7PL  
Plate: 06 row: A column: 09  
Seq primer: SKPL  
High quality sequence stop: 344.  
Location/Qualifiers

FEATURES  
source  
1..477  
/organism="Ascaris suum"  
/mol\_type="mRNA"  
/db\_xref="taxon:6253"  
/clone="As\_L3\_06A09"  
/sex="mixed"  
/dev\_stage="L3"  
/clone\_lib="Ascaris suum (parasitic nematode) L3 stage"  
/note="Vector: Lambda ZAP II, Site 1: EcoRI (5' end); The  
Site 2: (3' end); Ascaris suum is a parasite of pigs. The  
library was constructed from mRNA from Ascaris suum L3  
stage."

ORIGIN  
Query Match 83.3%; Score 15; DB 4; Length 477;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 CACGCGACACACTTC 18  
406 CACGCGACACACTTC 392

RESULT 14  
BU067587 481 bp mRNA linear EST 27-AUG-2002  
LOCUS  
DEFINITION  
1626\_F12\_K24ZT5 Mature perithecium Gibberella zeae cDNA, mRNA  
SEQUENCE  
BU067587  
BU067587.1 GI:22508776  
EST.  
Gibberella zeae  
Gibberella zeae  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
1 (bases 1 to 481)  
Trail, F., Xu, J.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.  
Analysis of expressed sequence tags from Gibberella zeae (anamorph  
Fusarium graminearum)  
Fungal Genet. Biol. 38 (2), 187-197 (2003)  
PUBMED  
22508120  
12620255  
COMMENT  
Contact: Frances Trail  
Department of Plant Biology  
Michigan State University  
East Lansing, MI 48824, USA  
Tel: 517 432 2939  
Fax: 517 353 1926  
Email: trail@msu.edu.  
Location/Qualifiers

FEATURES  
source  
1..481  
/organism="Gibberella zeae"  
/mol\_type="mRNA"  
/strain="NRRL 31084"

ORIGIN  
Query Match 83.3%; Score 15; DB 5; Length 481;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 CACGCGACACACTTC 18  
301 CACGCGACACACTTC 315

RESULT 15  
CA953980 483 bp mRNA linear EST 02-JAN-2003  
LOCUS  
DEFINITION  
K137a10.y1 Ascaris suum embryo SL1 TOPO v1 Ascaris suum cDNA 5'  
similar to SW:RL28\_MOUSE P41105 608 RIBOSOMAL PROTEIN L28.1, mRNA  
sequence.  
CA953980  
CA953980.1 GI:27470340  
EST.  
Ascaris suum (pig roundworm)  
Ascaris suum  
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida;  
Ascaridoidea; Ascarididae; Ascaris.  
1 (bases 1 to 483)  
McCarter, J., Clifton, S., Chapelli, B., Page, D., Martin, J.,  
Wyllie, T., Dente, M., Maria, M., Hillier, L., Kucaba, T., Theising, B.,  
Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Franklin, C.,  
Tsagarelashvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,  
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,  
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCam, R., Waterston, R. and Wilson, R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
The library was constructed by Claire Murphy and Dr. James McCarter  
at Washington University, St. Louis. Oligo (dT)-SL1 PCR based  
library. Embryo cDNA PCR products of size >400 nucleotides  
containing SL1 on the 5' end and oligo(dT) on the 3' end were  
non-directionally cloned into pCRIT-TOPO (Invitrogen) following the  
TOPO TA cloning protocol. 30-60 cell embryo material was provided  
by Dr. Richard Davis of City University of New York Graduate  
Center, College of Staten Island, Staten Island, NY  
(redavis@postbox.csi.cuny.edu).  
Putative full length read  
The vector to vector length is 588  
Seq primer: SL1 primer.  
Location/Qualifiers

FEATURES  
source  
1..483  
/organism="Ascaris suum"  
/mol\_type="mRNA"  
/db\_xref="taxon:6253"  
/dev\_stage="30-60 cell embryo"  
/lab\_host="DH10B"  
/clone\_lib="Ascaris suum embryo SL1 TOPO v1"  
/note="Vector: pCRIT-TOPO (Invitrogen); Site 1: EcoRI;  
Site 2: EcoRI; The library was constructed by Claire  
Murphy and Dr. James McCarter at Washington University,  
St. Louis. Oligo(dT)-SL1 PCR based library. Embryo cDNA  
PCR products of size >400 nucleotides containing SL1 on  
the 5' end and oligo(dT) on the 3' end were  
non-directionally cloned into pCRIT-TOPO (Invitrogen)  
following the TOPO TA cloning protocol. 30-60 cell embryo  
material was provided by Dr. Richard Davis of City



University of New York Graduate Center, College of Staten  
Island, Staten Island, NY (redavis@postbox.csi.cuny.edu). "

ORIGIN

Query Match 83.3%; Score 15; DB 6; Length 483;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACTTC 18  
|||||  
DB 421 CACGCGACACTTC 407

Search completed: February 10, 2005, 17:01:56  
Job time : 124.68 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41; Search time 29.935 Seconds

(without alignments)  
3460.943 Million cell updates/sec

Title: US-10-790-430-10

Sequence: 1 taccacgcgcacacattc 18

Scoring table: OLIGO NTC  
Gapop 60.0, Gapext 60.0

Searched: 4313806 seqs, 287787103 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

## Database:

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length  | DB ID | Description          |
|------------|-------|-------------|---------|-------|----------------------|
| 1          | 18    | 100.0       | 18      | 9     | US-09-872-051-10     |
| 2          | 18    | 100.0       | 18      | 9     | US-10-790-430-10     |
| 3          | 18    | 100.0       | 1183    | 9     | US-09-872-051-8      |
| 4          | 18    | 100.0       | 1183    | 18    | US-10-790-430-8      |
| 5          | 17    | 94.4        | 466     | 18    | US-10-425-115-114310 |
| 6          | 16    | 88.9        | 1338    | 17    | US-10-369-493-33178  |
| 7          | 15    | 83.3        | 687     | 18    | US-10-653-047-1012   |
| 8          | 15    | 83.3        | 721     | 18    | US-10-425-115-14852  |
| 9          | 15    | 83.3        | 747     | 18    | US-10-781-014-753    |
| 10         | 15    | 83.3        | 3309400 | 9     | US-09-738-626-1      |
| 11         | 14    | 77.8        | 435     | 11    | US-09-864-408A-1183  |

|      |    |      |      |    |                      |                    |
|------|----|------|------|----|----------------------|--------------------|
| C 12 | 14 | 77.8 | 435  | 18 | US-10-357-930-31761  | Sequence 31761, A  |
| C 13 | 14 | 77.8 | 435  | 18 | US-10-357-930-40730  | Sequence 40730, A  |
| C 14 | 14 | 77.8 | 437  | 18 | US-10-357-930-1395   | Sequence 1395, Ap  |
| C 15 | 14 | 77.8 | 484  | 18 | US-10-357-930-10564  | Sequence 10564, A  |
| C 16 | 14 | 77.8 | 732  | 18 | US-10-363-345A-1539  | Sequence 1539, Ap  |
| C 17 | 14 | 77.8 | 732  | 18 | US-10-363-345A-1540  | Sequence 1540, Ap  |
| C 18 | 13 | 72.2 | 314  | 9  | US-09-867-550-279    | Sequence 279, App  |
| C 19 | 13 | 72.2 | 419  | 17 | US-10-369-493-30016  | Sequence 30016, Ap |
| C 20 | 13 | 72.2 | 505  | 18 | US-10-425-115-114679 | Sequence 114679, A |
| C 21 | 13 | 72.2 | 564  | 18 | US-10-425-115-150679 | Sequence 150679, A |
| C 22 | 13 | 72.2 | 583  | 18 | US-10-425-115-75209  | Sequence 75209, A  |
| C 23 | 13 | 72.2 | 729  | 18 | US-10-363-345A-31705 | Sequence 31705, A  |
| C 24 | 13 | 72.2 | 729  | 18 | US-10-363-345A-31706 | Sequence 31706, A  |
| C 25 | 13 | 72.2 | 789  | 17 | US-10-424-599-79579  | Sequence 79579, A  |
| C 26 | 13 | 72.2 | 830  | 17 | US-10-425-114-33981  | Sequence 33981, A  |
| C 27 | 13 | 72.2 | 840  | 18 | US-10-363-345A-22085 | Sequence 22085, A  |
| C 28 | 13 | 72.2 | 840  | 18 | US-10-363-345A-22086 | Sequence 22086, A  |
| C 29 | 13 | 72.2 | 863  | 18 | US-10-425-115-6466   | Sequence 6466, Ap  |
| C 30 | 13 | 72.2 | 904  | 18 | US-10-363-345A-22235 | Sequence 22235, A  |
| C 31 | 13 | 72.2 | 904  | 18 | US-10-363-345A-22236 | Sequence 22236, A  |
| C 32 | 13 | 72.2 | 1023 | 18 | US-10-425-115-145387 | Sequence 145387, A |
| C 33 | 13 | 72.2 | 1125 | 17 | US-10-424-599-8880   | Sequence 8880, Ap  |
| C 34 | 13 | 72.2 | 1189 | 18 | US-10-767-701-12212  | Sequence 12212, A  |
| C 35 | 13 | 72.2 | 1251 | 18 | US-10-437-963-43431  | Sequence 43431, A  |
| C 36 | 13 | 72.2 | 1344 | 17 | US-10-365-742-35     | Sequence 35, Ap1   |
| C 37 | 13 | 72.2 | 1396 | 18 | US-10-363-345A-40673 | Sequence 40673, A  |
| C 38 | 13 | 72.2 | 1396 | 18 | US-10-363-345A-40674 | Sequence 40674, A  |
| C 39 | 13 | 72.2 | 1620 | 17 | US-10-282-112A-14408 | Sequence 14408, A  |
| C 40 | 13 | 72.2 | 1687 | 17 | US-10-425-114-15589  | Sequence 15589, A  |
| C 41 | 13 | 72.2 | 1743 | 18 | US-10-739-930-1844   | Sequence 1844, Ap  |
| C 42 | 13 | 72.2 | 1832 | 17 | US-10-425-114-25898  | Sequence 25898, A  |
| C 43 | 13 | 72.2 | 2000 | 9  | US-09-938-842A-4958  | Sequence 4958, Ap  |
| C 44 | 13 | 72.2 | 2000 | 11 | US-09-938-842A-4958  | Sequence 4958, Ap  |
| C 45 | 13 | 72.2 | 2587 | 17 | US-10-424-599-57089  | Sequence 57089, A  |

## ALIGNMENTS

RESULT 1  
US-09-872-051-10  
Sequence 10, Application US/09872051  
Patent No. US20020013960A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto CO  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(mk603) and Composition and Methods for Det.  
FILE REFERENCE: 38-21 (52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 10  
LENGTH: 18  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
NAME/KEY: source  
LOCATION: (1) ..(18)  
OTHER INFORMATION: zea maize plastid DNA and vector DNA  
US-09-872-051-10  
Query Match 100.0%; Score 18; DB 9; Length 18;

Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18  
DB 1 TACCACGCGACACTTC 18

## RESULT 2

US-10-790-430-10  
Sequence 10, Application US/10790430  
Publication No. US20040139493A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
TITLE OF INVENTION: Thereof  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/10/790,430  
CURRENT FILING DATE: 2004-03-01  
PRIOR APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 10  
LENGTH: 18  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(18)  
OTHER INFORMATION: zea maize plasmid DNA and vector DNA  
US-10-790-430-10

Query Match 100.0%; Score 18; DB 18; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18  
DB 1 TACCACGCGACACTTC 18

## RESULT 3

US-09-872-051-8  
Sequence 8, Application US/09872051  
Patent No. US20020013960A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
TITLE OF INVENTION: Thereof  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 1183  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(1183)  
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
OTHER INFORMATION: 165-381 construct vector DNA  
OTHER INFORMATION: 382-686 Zea maize plasmid genes, rpa1 and rpaA  
OTHER INFORMATION: 687-1183 Zea maize genomic DNA  
US-09-872-051-8

Query Match 100.0%; Score 18; DB 9; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18  
DB 373 TACCACGCGACACTTC 390

## RESULT 4

US-10-790-430-8  
Sequence 8, Application US/10790430  
Publication No. US20040139493A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
TITLE OF INVENTION: Thereof  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/10/790,430  
CURRENT FILING DATE: 2004-03-01  
PRIOR APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 1183  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(1183)  
OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
OTHER INFORMATION: 165-381 construct vector DNA  
OTHER INFORMATION: 382-686 Zea maize plasmid genes, rpa1 and rpaA  
OTHER INFORMATION: 687-1183 Zea maize genomic DNA  
US-10-790-430-8

Query Match 100.0%; Score 18; DB 18; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGCGACACTTC 18  
DB 373 TACCACGCGACACTTC 390

## RESULT 5

US-10-425-115-114310

```
/ Sequence 114310, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ PRIOR FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 114310
/ LENGTH: 466
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_35734C.1
US-10-425-115-114310

Query Match          94.4%; Score 17; DB 18; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TACCACGCGACACACTT 17
Db      267 TACCACGCGACACACTT 283

RESULT 6
US-10-369-493-33178/c
/ Sequence 33178, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 33178
/ LENGTH: 1338
/ TYPE: DNA
/ ORGANISM: Xylella fastidiosa
US-10-369-493-33178

Query Match          88.9%; Score 16; DB 17; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ACCACGCGACACACTT 17
Db      513 ACCACGCGACACACTT 498

RESULT 7
US-10-653-047-1012
/ Sequence 1012, Application US/10653047
/ Publication No. US20040229367A1
/ GENERAL INFORMATION:
/ APPLICANT: Randy M. Berka
/ APPLICANT: Michael W. Rey
/ APPLICANT: Jeffrey R. Shuster
/ APPLICANT: Sakari Kauppinen
/ APPLICANT: Ib Groch Clausen
```

```
/ APPLICANT: Peter Bjarke Olsen
/ TITLE OF INVENTION: Methods For Monitoring Multiple Gene
/ FILE REFERENCE: 5849,200-US
/ CURRENT APPLICATION NUMBER: US/10/653,047
/ CURRENT FILING DATE: 2003-08-29
/ PRIOR APPLICATION NUMBER: US/09/533,559
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 09/273,623
/ PRIOR FILING DATE: 1999-03-22
/ NUMBER OF SEQ ID NOS: 7860
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1012
/ LENGTH: 687
/ TYPE: DNA
/ ORGANISM: Fusarium venenatum
US-10-653-047-1012

Query Match          83.3%; Score 15; DB 18; Length 687;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 CACGCGACACACTTC 18
Db      531 CACGCGACACACTTC 545

RESULT 8
US-10-425-115-14852
/ Sequence 14852, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 14852
/ LENGTH: 721
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)-(721)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_113531C.1
US-10-425-115-14852

Query Match          83.3%; Score 15; DB 18; Length 721;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 CACGCGACACACTTC 18
Db      476 CACGCGACACACTTC 490

RESULT 9
US-10-781-014-753/c
/ Sequence 753, Application US/10781014
/ Publication No. US20040180408A1
/ GENERAL INFORMATION:
/ APPLICANT: Pompejus, Markus
/ APPLICANT: Kroger, Burkhard
/ APPLICANT: Schröder, Hartwig
/ APPLICANT: Zeidler, Oskar
/ APPLICANT: Haberer, Gregor
```

```
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
/ TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
/ TITLE OF INVENTION: PRODUCTION
/ FILE REFERENCE: BGI-126CCPN
/ CURRENT APPLICATION NUMBER: US/10/781,014
/ PRIOR FILING DATE: 2004-02-17
/ PRIOR APPLICATION NUMBER: US 09/602,740
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141,031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/143,208
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 60/151,572
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19931412.8
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931413.6
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931419.5
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931420.9
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931424.1
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931428.4
/ PRIOR FILING DATE: 1999-07-08
/ REMAINING PRIOR APPLICATION data removed - See file Wrapper or PLM.
/ NUMBER OF SEQ ID NOS: 784
/ SEQ ID NO: 753
/ LENGTH: 747
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (101)..(724)
/ OTHER INFORMATION: RXN02554
/ US-10-781-014-753

Query Match      83.3%; Score 15; DB 18; Length 747;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 CACGCGACACTTC 18
DB      21 CACGCGACACTTC 7
```

```
RESULT 10
US-09-738-626-1/c
/ Sequence 1, Application US/09738626
/ Publication No. US20020197605A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIKO
/ APPLICANT: HAYASHI, MIKIRO
/ APPLICANT: OCHIAI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAOKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ PRIOR FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
```

```
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO: 1
/ LENGTH: 3309400
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ US-09-738-626-1
```

```
Query Match      83.3%; Score 15; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 CACGCGACACTTC 18
DB      3028188 CACGCGACACTTC 3028174
```

```
RESULT 11
US-09-864-408A-1183
/ Sequence 1183, Application US/09864408A
/ Publication No. US20040009474A1
/ GENERAL INFORMATION:
/ APPLICANT: Leach, Martin D.
/ APPLICANT: Shinkets, Richard A.
/ TITLE OF INVENTION: NO. US20040009474A1 Human Polynucleotides and Polypeptides Enc
/ FILE REFERENCE: 21402-012
/ CURRENT APPLICATION NUMBER: US/09/864,408A
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 60/206,690
/ PRIOR FILING DATE: 2000-05-24
/ NUMBER OF SEQ ID NOS: 9068
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 1183
/ LENGTH: 435
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-864-408A-1183
```

```
Query Match      77.8%; Score 14; DB 11; Length 435;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 ACGCGACACTTC 18
DB      261 ACGCGACACTTC 274
```

```
RESULT 12
US-10-357-930-31761/c
/ Sequence 31761, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-0078CN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
```

;; PRIOR FILING DATE: 2000-12-13  
;; NUMBER OF SEQ ID NOS: 62232  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 31761  
;; LENGTH: 435  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-357-930-31761

Query Match 77.8%; Score 14; DB 18; Length 435;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTT 17  
Db 286 CACGCGACACACTT 273

RESULT 13  
US-10-357-930-40730/c  
; Sequence 40730, Application US/10357930  
; Publication No. US20040259086A1

;; GENERAL INFORMATION:  
;; APPLICANT: Schlegel, Robert  
;; APPLICANT: Endege, Wilson  
;; APPLICANT: Monahan, John  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
;; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
;; FILE REFERENCE: MRI-007BCN  
;; CURRENT APPLICATION NUMBER: US/10/357,930  
;; CURRENT FILING DATE: 2003-02-04

;; PRIOR APPLICATION NUMBER: 09/785,276  
;; PRIOR FILING DATE: 2003-02-16  
;; PRIOR APPLICATION NUMBER: 60/183,319  
;; PRIOR FILING DATE: 2000-02-17  
;; PRIOR APPLICATION NUMBER: 60/189,862  
;; PRIOR FILING DATE: 2000-03-16  
;; PRIOR APPLICATION NUMBER: 60/207,454  
;; PRIOR FILING DATE: 2000-05-25  
;; PRIOR APPLICATION NUMBER: 60/211,314  
;; PRIOR FILING DATE: 2000-06-09  
;; PRIOR APPLICATION NUMBER: 60/219,007  
;; PRIOR FILING DATE: 2000-07-18  
;; PRIOR APPLICATION NUMBER: 60/255,281  
;; PRIOR FILING DATE: 2000-12-13  
;; NUMBER OF SEQ ID NOS: 62232  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 40730  
;; LENGTH: 435  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-357-930-40730

Query Match 77.8%; Score 14; DB 18; Length 435;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTT 17  
Db 286 CACGCGACACACTT 273

RESULT 14  
US-10-357-930-1395/c  
; Sequence 1395, Application US/10357930  
; Publication No. US20040259086A1

;; GENERAL INFORMATION:  
;; APPLICANT: Schlegel, Robert  
;; APPLICANT: Endege, Wilson  
;; APPLICANT: Monahan, John  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF

;; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
;; FILE REFERENCE: MRI-007BCN  
;; CURRENT APPLICATION NUMBER: US/10/357,930  
;; CURRENT FILING DATE: 2003-02-04

;; PRIOR APPLICATION NUMBER: 09/785,276  
;; PRIOR FILING DATE: 2003-02-16  
;; PRIOR APPLICATION NUMBER: 60/183,319  
;; PRIOR FILING DATE: 2000-02-17  
;; PRIOR APPLICATION NUMBER: 60/189,862  
;; PRIOR FILING DATE: 2000-03-16  
;; PRIOR APPLICATION NUMBER: 60/207,454  
;; PRIOR FILING DATE: 2000-05-25  
;; PRIOR APPLICATION NUMBER: 60/211,314  
;; PRIOR FILING DATE: 2000-06-09  
;; PRIOR APPLICATION NUMBER: 60/219,007  
;; PRIOR FILING DATE: 2000-07-18  
;; PRIOR APPLICATION NUMBER: 60/255,281  
;; PRIOR FILING DATE: 2000-12-13  
;; NUMBER OF SEQ ID NOS: 62232  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1395  
;; LENGTH: 437  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens

;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 36, 86, 87  
;; OTHER INFORMATION: n = A,T,C or G  
US-10-357-930-1395

Query Match 77.8%; Score 14; DB 18; Length 437;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTT 17  
Db 288 CACGCGACACACTT 275

RESULT 15  
US-10-357-930-10564/c  
; Sequence 10564, Application US/10357930  
; Publication No. US20040259086A1

;; GENERAL INFORMATION:  
;; APPLICANT: Schlegel, Robert  
;; APPLICANT: Endege, Wilson  
;; APPLICANT: Monahan, John  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
;; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
;; FILE REFERENCE: MRI-007BCN  
;; CURRENT APPLICATION NUMBER: US/10/357,930  
;; CURRENT FILING DATE: 2003-02-04  
;; PRIOR APPLICATION NUMBER: 09/785,276  
;; PRIOR FILING DATE: 2003-02-16  
;; PRIOR APPLICATION NUMBER: 60/183,319  
;; PRIOR FILING DATE: 2000-02-17  
;; PRIOR APPLICATION NUMBER: 60/189,862  
;; PRIOR FILING DATE: 2000-03-16  
;; PRIOR APPLICATION NUMBER: 60/207,454  
;; PRIOR FILING DATE: 2000-05-25  
;; PRIOR APPLICATION NUMBER: 60/211,314  
;; PRIOR FILING DATE: 2000-06-09  
;; PRIOR APPLICATION NUMBER: 60/219,007  
;; PRIOR FILING DATE: 2000-07-18  
;; PRIOR APPLICATION NUMBER: 60/255,281  
;; PRIOR FILING DATE: 2000-12-13  
;; NUMBER OF SEQ ID NOS: 62232  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 10564  
;; LENGTH: 484  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens

US-10-357-930-10564

Query Match 77.8%; Score 14; DB 18; Length 484;

Best Local Similarity 100.0%; Pred.No. 28;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CACGCGACACACTT 17  
 |||||  
 Db 335 CACGCGACACACTT 322

Search completed: February 9, 2005, 11:12:45  
 Job time : 34.935 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 31.5981 Seconds  
(without alignments)  
3460.943 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19  
Sequence: 1 tctagcgcgcacgcgtcg 19

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description         |
|------------|-------|-------------|---------|----|---------------------|
| 1          | 19    | 100.0       | 19      | 9  | US-09-872-051-9     |
| 2          | 19    | 100.0       | 19      | 18 | US-10-790-430-9     |
| 3          | 19    | 100.0       | 498     | 9  | US-09-872-051-7     |
| 4          | 19    | 100.0       | 498     | 18 | US-10-790-430-7     |
| 5          | 16    | 84.2        | 1044    | 15 | US-10-156-761-1684  |
| 6          | 16    | 84.2        | 35167   | 18 | US-10-645-794-5     |
| 7          | 16    | 84.2        | 35167   | 18 | US-10-645-883A-1    |
| 8          | 16    | 84.2        | 9025608 | 15 | US-10-156-761-1     |
| 9          | 15    | 78.9        | 521     | 17 | US-10-424-599-43339 |
| 10         | 15    | 78.9        | 797     | 18 | US-10-767-701-9887  |
| 11         | 15    | 78.9        | 1339    | 18 | US-10-437-963-26586 |

|    |    |      |       |    |                      |                    |
|----|----|------|-------|----|----------------------|--------------------|
| 12 | 15 | 78.9 | 1398  | 18 | US-10-437-963-33098  | Sequence 33098, A  |
| 13 | 15 | 78.9 | 1780  | 18 | US-10-473-390-1      | Sequence 1, Appl1  |
| 14 | 15 | 78.9 | 5013  | 18 | US-10-437-963-60106  | Sequence 60106, A  |
| 15 | 15 | 78.9 | 14800 | 18 | US-10-473-380-3      | Sequence 3, Appl1  |
| 16 | 14 | 73.7 | 218   | 18 | US-10-425-115-172757 | Sequence 172757, A |
| 17 | 14 | 73.7 | 226   | 18 | US-10-425-115-47913  | Sequence 47913, A  |
| 18 | 14 | 73.7 | 228   | 9  | US-09-983-965-4908   | Sequence 4908, Ap  |
| 19 | 14 | 73.7 | 228   | 17 | US-10-369-493-32240  | Sequence 32240, A  |
| 20 | 14 | 73.7 | 292   | 9  | US-09-960-352-7737   | Sequence 7737, Ap  |
| 21 | 14 | 73.7 | 328   | 17 | US-10-424-599-79875  | Sequence 79875, A  |
| 22 | 14 | 73.7 | 328   | 17 | US-10-425-115-174641 | Sequence 174641, A |
| 23 | 14 | 73.7 | 414   | 10 | US-09-918-995-17011  | Sequence 17011, A  |
| 24 | 14 | 73.7 | 436   | 9  | US-09-983-965-2307   | Sequence 2307, Ap  |
| 25 | 14 | 73.7 | 465   | 16 | US-10-029-386-25290  | Sequence 25290, A  |
| 26 | 14 | 73.7 | 473   | 18 | US-10-425-115-29397  | Sequence 29397, A  |
| 27 | 14 | 73.7 | 500   | 9  | US-09-754-997A-43    | Sequence 43, Appl  |
| 28 | 14 | 73.7 | 512   | 17 | US-10-424-599-17151  | Sequence 17151, A  |
| 29 | 14 | 73.7 | 549   | 18 | US-10-437-963-25717  | Sequence 25717, A  |
| 30 | 14 | 73.7 | 610   | 18 | US-10-767-701-23177  | Sequence 23177, A  |
| 31 | 14 | 73.7 | 619   | 18 | US-10-767-795-5965   | Sequence 5965, Ap  |
| 32 | 14 | 73.7 | 619   | 18 | US-10-425-115-55360  | Sequence 55360, A  |
| 33 | 14 | 73.7 | 629   | 18 | US-10-437-963-61015  | Sequence 61015, A  |
| 34 | 14 | 73.7 | 635   | 18 | US-10-425-115-164400 | Sequence 164400, A |
| 35 | 14 | 73.7 | 642   | 18 | US-10-437-963-9105   | Sequence 9105, Ap  |
| 36 | 14 | 73.7 | 676   | 18 | US-10-437-963-90735  | Sequence 90735, A  |
| 37 | 14 | 73.7 | 754   | 18 | US-10-437-963-85113  | Sequence 85113, A  |
| 38 | 14 | 73.7 | 834   | 17 | US-10-260-238-824    | Sequence 824, App  |
| 39 | 14 | 73.7 | 848   | 18 | US-10-425-115-176769 | Sequence 176769, A |
| 40 | 14 | 73.7 | 870   | 17 | US-10-260-238-4711   | Sequence 4711, Ap  |
| 41 | 14 | 73.7 | 960   | 17 | US-10-424-599-127149 | Sequence 127149, A |
| 42 | 14 | 73.7 | 1146  | 17 | US-10-282-122A-14674 | Sequence 14674, A  |
| 43 | 14 | 73.7 | 1209  | 10 | US-09-826-509-556    | Sequence 556, App  |
| 44 | 14 | 73.7 | 1209  | 18 | US-10-472-362-2      | Sequence 2, Appl1  |
| 45 | 14 | 73.7 | 1209  | 19 | US-10-925-095-556    | Sequence 556, App  |

#### ALIGNMENTS

RESULT 1  
US-09-872-051-9  
; Sequence 9, Application US/09872051  
; Patent No. US20020013960A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OR INVENTION: Corn Event PV-ZMGN32 (nke603) and Composition and Methods for Det  
; FILE REFERENCE: 38-21 (52258)B  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; NAME/KEY: source  
; LOCATION: (1)..(19)  
; OTHER INFORMATION: Zea maize genomic and vector DNA  
US-09-872-051-9  
Query Match 100.0%; Score 19; DB 9; Length 19;

Best Local Similarity 100.0%; Pred. No. 0.041;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19  
Db 1 TGTAGCGGCCACGCGTGG 19

## RESULT 2

US-10-790-430-9  
Sequence 9, Application US/10790430  
Publication No. US20040139493A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/10/790,430  
CURRENT FILING DATE: 2004-03-01  
PRIOR APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 19  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(19)  
OTHER INFORMATION: Zea maize genomic and vector DNA  
US-10-790-430-9

Query Match 100.0%; Score 19; DB 18; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.041;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19  
Db 1 TGTAGCGGCCACGCGTGG 19

## RESULT 3

US-09-872-051-7  
Sequence 7, Application US/09872051  
Patent No. US20020013960A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 7  
LENGTH: 498  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(498)  
OTHER INFORMATION: 1-304 Zea maize genomic DNA  
OTHER INFORMATION: 305-349 construct vector DNA  
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA  
US-09-872-051-7

Query Match 100.0%; Score 19; DB 9; Length 498;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19  
Db 296 TGTAGCGGCCACGCGTGG 314

## RESULT 4

US-10-790-430-7  
Sequence 7, Application US/10790430  
Publication No. US20040139493A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/10/790,430  
CURRENT FILING DATE: 2004-03-01  
PRIOR APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 7  
LENGTH: 498  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(498)  
OTHER INFORMATION: 1-304 Zea maize genomic DNA  
OTHER INFORMATION: 305-349 construct vector DNA  
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA  
US-10-790-430-7

Query Match 100.0%; Score 19; DB 18; Length 498;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCACGCGTGG 19  
Db 296 TGTAGCGGCCACGCGTGG 314

## RESULT 5

US-10-156-761-1684/c  
Sequence 1684, Application US/10156761  
Publication No. US20030119018A1

```

; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1684
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1044)
; US-10-156-761-1684

```

```

Query Match      84.2%; Score 16; DB 15; Length 1044;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 TAGCGGCCCGCGGTG 18
Db      359 TAGCGGCCCGCGGTG 344

```

```

RESULT 6
US-10-645-794-5/c
; Sequence 5, Application US/10645794
; Publication No. US20040106194A1
; GENERAL INFORMATION:
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chastain, Michael
; APPLICANT: Sandig, Volker
; APPLICANT: Emml, Emilio A.
; APPLICANT: Shiver, John W.
; APPLICANT: Casimiro, Danilo R.
; APPLICANT: Kaslow, David C.
; APPLICANT: Morsy, Manal
; TITLE OF INVENTION: METHODS FOR PROPAGATING ADENOVIRUS AND
; FILE REFERENCE: 20699Y
; CURRENT APPLICATION NUMBER: US/10/645,794
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: 60/458,825
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 60/455,312
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/455,234
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/405,182
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 35167
; TYPE: DNA
; ORGANISM: adenovirus serotype 24
; US-10-645-794-5

```

```

Query Match      84.2%; Score 16; DB 18; Length 35167;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGTAGCGGCCCGCGG 16
Db      12091 TGTAGCGGCCCGCGG 12076

```

```

RESULT 7
US-10-645-883A-1/c
; Sequence 1, Application US/10645883A
; Publication No. US2004018555A1
; GENERAL INFORMATION:
; APPLICANT: Emml, Emilio A.
; APPLICANT: Shiver, John W.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Casimiro, Danilo R.
; APPLICANT: Chastain, Michael
; APPLICANT: Kaslow, David C.
; APPLICANT: Morsy, Manal
; TITLE OF INVENTION: ADENOVIRUS SEROTYPE 24 VECTORS, NUCLEIC
; FILE REFERENCE: 21366
; CURRENT APPLICATION NUMBER: US/10/645,883A
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: 60/455,312
; PRIOR FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 35167
; TYPE: DNA
; ORGANISM: adenovirus serotype 24
; US-10-645-883A-1

```

```

Query Match      84.2%; Score 16; DB 18; Length 35167;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGTAGCGGCCCGCGG 16
Db      12091 TGTAGCGGCCCGCGG 12076

```

```

RESULT 8
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
; US-10-156-761-1

```

```

Query Match      84.2%; Score 16; DB 15; Length 9025608;

```

Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCACGCGTG 18  
|||||  
DB 2073859 TAGCGGCCACGCGTG 2073874

## RESULT 9

US-10-424-599-43339/c  
; Sequence 43339, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 43339  
; LENGTH: 521  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_139133C.1  
US-10-424-599-43339

Query Match 78.9%; Score 15; DB 17; Length 521;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TAGCGGCCACGCGT 17  
|||||  
DB 195 TAGCGGCCACGCGT 181

RESULT 10  
US-10-767-701-9887/c  
; Sequence 9887, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 9887  
; LENGTH: 797  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAV03-CIUS39666\_1  
US-10-767-701-9887

Query Match 78.9%; Score 15; DB 18; Length 797;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCACGCG 16  
|||||  
DB 195 GTAGCGGCCACGCG 181

RESULT 11  
US-10-437-963-26588/c

Sequence 26588, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 26588  
; LENGTH: 1339  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_31363C.1  
US-10-437-963-26588

Query Match 78.9%; Score 15; DB 18; Length 1339;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCACGCGTG 18  
|||||  
DB 121 AGCGGCCACGCGTG 107

RESULT 12  
US-10-437-963-33098/c  
; Sequence 33098, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 33098  
; LENGTH: 1398  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_37243C.1  
US-10-437-963-33098

Query Match 78.9%; Score 15; DB 18; Length 1398;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCACGCGTG 18  
|||||  
DB 617 AGCGGCCACGCGTG 603

RESULT 13  
US-10-473-390-1/c  
; Sequence 1, Application US/10473390

```
; Publication No. US20040248112A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001206-US
; CURRENT APPLICATION NUMBER: US/10/473,390
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/US02/09743
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,923
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1780
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-390-1
```

```
Query Match          78.9%; Score 15; DB 18; Length 1780;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 AGCGGCCCAAGCGGTG 18
Db      51 AGCGGCCCAAGCGGTG 37
```

```
RESULT 14
US-10-437-963-60106
; Sequence 60106, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barzak, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 60106
; LENGTH: 5013
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61666C.1
US-10-437-963-60106
```

```
Query Match          78.9%; Score 15; DB 18; Length 5013;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 GTAGCGGCCCAAGCGG 16
Db      2857 GTAGCGGCCCAAGCGG 2871
```

```
RESULT 15
US-10-473-390-3/c
; Sequence 3, Application US/10473390
; Publication No. US20040248112A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
```

```
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001206-US
; CURRENT APPLICATION NUMBER: US/10/473,390
; PRIOR FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: PCT/US02/09743
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,923
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-390-3
```

```
Query Match          78.9%; Score 15; DB 18; Length 14800;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 AGCGGCCCAAGCGGTG 18
Db      2050 AGCGGCCCAAGCGGTG 2036
```

```
Search completed: February 9, 2005, 11:12:40
Job time : 43.5981 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 18.6055 Seconds  
(without alignments)  
5727.097 Million cell updates/sec

Title: US-10-790-430-10

Sequence: 18  
1 taccacgcgcacacttc 18

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 18    | 100.0       | 18     | 6     | ABK15247    |
| 2          | 18    | 100.0       | 18     | 6     | ABK15245    |
| 3          | 16    | 88.9        | 1338   | 13    | ADSS7504    |
| 4          | 15    | 83.3        | 687    | 3     | AAF08489    |
| 5          | 15    | 83.3        | 747    | 4     | AAE71736    |
| 6          | 15    | 83.3        | 309400 | 5     | AAH68534    |
| 7          | 15    | 83.3        | 349980 | 5     | AAH68533    |
| 8          | 14    | 77.8        | 320    | 13    | ADQ54932    |
| 9          | 14    | 77.8        | 435    | 5     | ABV40711    |
| 10         | 14    | 77.8        | 435    | 5     | ABV31743    |
| 11         | 14    | 77.8        | 435    | 5     | ABN75645    |
| 12         | 14    | 77.8        | 437    | 5     | ABV01404    |
| 13         | 14    | 77.8        | 484    | 5     | ABV10573    |
| 14         | 14    | 77.8        | 732    | 6     | ABQ14948    |
| 15         | 14    | 77.8        | 732    | 6     | ABQ14949    |
| 16         | 14    | 77.8        | 2000   | 6     | ADA72018    |
| 17         | 14    | 77.8        | 5813   | 4     | ABL07158    |
| 18         | 13    | 72.2        | 233    | 2     | AAV20257    |
| 19         | 13    | 72.2        | 279    | 6     | ABN16209    |
| 20         | 13    | 72.2        | 296    | 10    | ABX85992    |

# ALIGNMENTS

|    |    |      |       |    |          |                    |
|----|----|------|-------|----|----------|--------------------|
| 21 | 13 | 72.2 | 314   | 6  | ABQ98333 | Abq98333 Human ORF |
| 22 | 13 | 72.2 | 419   | 13 | ADSS4342 | Adss4342 Bacterial |
| 23 | 13 | 72.2 | 729   | 6  | ABQ45115 | Abq45115 Oligonuc  |
| 24 | 13 | 72.2 | 729   | 6  | ABQ45114 | Abq45114 Oligonuc  |
| 25 | 13 | 72.2 | 840   | 6  | ABQ35494 | Abq35494 Oligonuc  |
| 26 | 13 | 72.2 | 840   | 6  | ABQ35495 | Abq35495 Oligonuc  |
| 27 | 13 | 72.2 | 904   | 6  | ABQ35644 | Abq35644 Oligonuc  |
| 28 | 13 | 72.2 | 904   | 6  | ABQ35645 | Abq35645 Oligonuc  |
| 29 | 13 | 72.2 | 1344  | 10 | ADG73042 | Adg73042 DNA encod |
| 30 | 13 | 72.2 | 1344  | 10 | ADL12097 | Adl12097 Pseudomon |
| 31 | 13 | 72.2 | 1396  | 6  | ABQ54082 | Abq54082 Oligonuc  |
| 32 | 13 | 72.2 | 1396  | 6  | ABQ54083 | Abq54083 Oligonuc  |
| 33 | 13 | 72.2 | 1467  | 5  | AA888878 | AA888878 DNA encod |
| 34 | 13 | 72.2 | 1620  | 8  | ACA26538 | ACA26538 Prokaryot |
| 35 | 13 | 72.2 | 2000  | 6  | ABZ17153 | Abz17153 Arabidops |
| 36 | 13 | 72.2 | 2761  | 4  | ABL04601 | Ab104601 Drosophi  |
| 37 | 13 | 72.2 | 3090  | 4  | ABL15163 | Ab115163 Drosophi  |
| 38 | 13 | 72.2 | 4682  | 4  | ABL04600 | Ab104600 Drosophi  |
| 39 | 13 | 72.2 | 5580  | 4  | ABL15162 | Ab115162 Drosophi  |
| 40 | 13 | 72.2 | 8246  | 6  | ABL13202 | Ab113202 Human imm |
| 41 | 13 | 72.2 | 9080  | 4  | ABL19130 | Ab119130 Human imm |
| 42 | 13 | 72.2 | 13875 | 4  | AA859616 | AA859616 Propionib |
| 43 | 13 | 72.2 | 13875 | 8  | ACE64545 | ACE64545 Propionib |
| 44 | 13 | 72.2 | 17369 | 4  | ABL02754 | Ab102754 Drosophi  |
| 45 | 13 | 72.2 | 28199 | 8  | ADA68658 | Ada68658 Takifugu  |

# RESULT 1

|    |  |                       |
|----|--|-----------------------|
| ID | ABK15247   | standard, DNA, 18 BP. |
| AC | ABK15247;  |                       |
| DT | 08-MAY-2002  | (first entry)         |
| DE | Corn transgene junction probe #2.  |                       |
| XX | Probe; ss; rice actin 1 promoter; RA1; RA1 intron;                       |                       |
| KW | chloroplast transit peptide gene; glyphosate resistance; corn;           |                       |
| KW | 5-enol-pyruvylphosphoketide-3-phosphatase synthase; EPSPS; Hsp70 intron; |                       |
| KW | transcriptional terminator; cauliflower mosaic virus 35S promoter;       |                       |
| KW | PV-ZWGT32; transgenic; nk603.  |                       |
| XX |  |                       |
| OS | Zea mays.  |                       |
| OS | Synthetic.   |                       |
| XX |  |                       |
| PN | EP167531-A1.   |                       |
| XX |  |                       |
| PD | 02-JAN-2002.   |                       |
| XX |  |                       |
| PF | 15-JUN-2001; 2001EP-00202314.  |                       |
| XX |  |                       |
| PR | 22-JUN-2000; 2000US-0213567P.  |                       |
| PR | 13-OCT-2000; 2000US-0240014P.  |                       |
| PR | 13-OCT-2000; 2000US-0241215P.  |                       |
| XX |  |                       |
| PA | (MONS ) MONSANTO TECHNOLOGY LLC.   |                       |
| XX |  |                       |
| FL | Behr CF, Hironaka C, Heck GR, You J;                                     |                       |
| XX |  |                       |
| DR | WPI; 2002-165871/22.   |                       |
| XX |  |                       |
| PT | Novel DNA construct useful for producing a corn plant that tolerates     |                       |
| PT | application of glyphosate herbicide, comprises two transgene expression  |                       |
| PT | cassettes.   |                       |
| XX |  |                       |
| PS | Claim 8; Page 3; 25pp; English.  |                       |
| XX |  |                       |
| CC | The invention relates to a DNA construct (1) comprising a first (F1) and |                       |
| CC | a second (F2) expression cassette. F1 of the DNA construct, in operable  |                       |

CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
CC PV-ZMGT32(nk603) of 498 or 1183bp, or primers and probes derived from  
CC them (used to detect transgene junction points). The construct is useful  
CC for producing a corn plant that tolerates application of glyphosate  
CC herbicide by transforming a corn cell with, selecting the corn cell for  
CC tolerance to application of glyphosate, and growing the corn cell into a  
CC fertile corn plant. The present sequence is a transgene junction probe  
CC sequence consisting of corn plasmid DNA and vector sequences  
XX  
SQ Sequence 18 BP; 5 A; 8 C; 2 G; 3 T; 0 U; 0 Other;  
Query Match 100.0%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. NO. 0.18; Mismatches 0; Gaps 0;  
Matches 18; Conservative 0; Indels 0; Gaps 0;  
QY 1 TACCACGCGACACACTTC 18  
Db 1 TACCACGCGACACACTTC 18  
RESULT 2  
ABK15245  
ID ABK15245 standard; DNA; 1183 BP.  
XX  
AC ABK15245;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Corn nk603 3' transgene/genomic sequence PCR fragment.  
XX  
KM de: rice actin 1 promoter; RA1; RA1 intron;  
KM chloroplast transit peptide gene; glyphosate resistance; corn;  
KM 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
KM transcriptional terminator; cauliflower mosaic virus 35S promoter;  
KM PV-ZMGT32; transgenic; nk603.  
XX  
OS Zea mays.  
OS Escherichia coli.  
OS Agrobacterium tumefaciens.  
OS Synthetic.  
OS Chimeric.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1..164  
FT /\*tag= a  
FT /note= "T-AGRTU.nos vector sequence"  
FT misc\_feature 165..381  
FT /\*tag= b  
FT /note= "Vector sequence"  
FT misc\_feature 382..686  
FT /\*tag= c  
FT /note= "Corn plasmid genes tps11/rpoA partial sequence"  
FT misc\_feature 687..1183  
FT /\*tag= d  
FT /note= "Corn genomic sequence"  
XX  
PN EP1167531-A1.  
XX  
PD 02-JAN-2002.  
XX  
PF 15-JUN-2001; 2001EP-00202314.  
XX  
PR 22-JUN-2000; 2000US-0213567P.  
PR 13-OCT-2000; 2000US-0240014P.  
PR 13-OCT-2000; 2000US-0241215P.  
XX  
PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
PI Behr CF, Hironaka C, Heck GR, You J;

XX  
DR WPI; 2002-165871/22.  
XX  
PT Novel DNA construct useful for producing a corn plant that tolerates  
PT application of glyphosate herbicide, comprises two transgene expression  
PT cassettes.  
XX  
PS Claim 3; Page 16-17; 25pp; English.  
XX  
CC The invention relates to a DNA construct (I) comprising a first (F1) and  
CC a second (F2) expression cassette. F1 of the DNA construct, in operable  
CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
CC PV-ZMGT32(nk603) of 498 or 1183bp, or primers and probes derived from  
CC them (used to detect transgene junction points). The construct is useful  
CC for producing a corn plant that tolerates application of glyphosate  
CC herbicide by transforming a corn cell with, selecting the corn cell for  
CC tolerance to application of glyphosate, and growing the corn cell into a  
CC fertile corn plant. The present sequence is the 1183bp amplicon  
CC consisting of the Agrobacterium transcriptional terminator sequence,  
CC vector sequences and corn sequences  
XX  
SQ Sequence 1183 BP; 295 A; 289 C; 280 G; 319 T; 0 U; 0 Other;  
Query Match 100.0%; Score 18; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. NO. 0.17; Mismatches 0; Gaps 0;  
Matches 18; Conservative 0; Indels 0; Gaps 0;  
QY 1 TACCACGCGACACACTTC 18  
Db 373 TACCACGCGACACACTTC 390  
RESULT 3  
ADS57504/C  
ID ADS57504 standard; cDNA; 1338 BP.  
XX  
AC ADS57504;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polynucleotide #9491.  
XX  
KM Recombinant DNA construct; transformed plant; improved plant property;  
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KM pathogen tolerance; pest tolerance; plant disease resistance;  
KM cell cycle pathway modification; plant growth regulator;  
KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KM bacterial polynucleotide; gene; ss.  
XX  
OS Bacteria.  
XX  
FN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI; 2004-061375/06.



XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
PS Claim 1, SEQ ID NO 33178; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomanan  
CC production. This sequence represents a bacterial polynucleotide used in  
CC the scope of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/Sequence.html](http://seqdata.uspto.gov/Sequence.html).  
XX

XX Sequence 1338 BP; 241 A; 281 C; 463 G; 353 T; 0 U; 0 Other;

XX Query Match 88.9%; Score 16; DB 13; Length 1338;

XX Best Local Similarity 100.0%; Pred. No. 2.9; Mismatches 0; Indels 0; Gaps 0;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Db 2 ACCACGGCAGACACTT 17  
513 ACCACGGCAGACACTT 498

XX RESULT 4

XX AAF08489 ID AAF08489 standard; cDNA; 687 BP.

XX AC AAF08489;

XX DT 13-MAR-2001 (first entry)

XX DE Fusarium venenatum EST SEQ ID NO:1012.

XX Multiple gene expression; filamentous fungal cell, EST;  
XX expressed sequence tag; Fusarium venenatum; *Aspergillus niger*;  
XX *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;  
XX culture condition; environmental stress; spore morphogenesis;  
XX metabolic pathway engineering; catabolic pathway engineering; ss.

XX OS Fusarium venenatum.

XX PN WO200056762-A2.

XX PD 28-SEP-2000.

XX PF 22-MAR-2000; 2000WO-US007781.

XX PR 22-MAR-1999; 99US-00273623.

XX PA (NOVO ) NOVO NORDISK BIOTECH INC.

XX RA (NOVO ) NOVO NORDISK AS.

XX PI Berka RM, Rey WM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

XX DR MPI; 2000-594572/56.

XX Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags.  
XX

XX Claim 86; Page 768; 3161pp; English.

CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring the  
CC global expression of genes from FF cells allows the production potential  
CC of the microorganisms to be improved. New genes may be discovered,  
CC possible functions of unknown open reading frames can be identified and  
CC gene copy number variation and stability can be monitored. The expression  
CC of genes can be used to study how FF cells adapt to changes in culture  
CC conditions, environmental stress, spore morphogenesis, recombination,  
CC metabolic or catabolic pathway engineering. Using ESTs provides several  
CC advantages over genomic or random cDNA clones including elimination of  
CC redundancy as one spot on an array equals one gene or open reading frame,  
CC and organization of the microarray based on function of the gene  
CC products to facilitate analysis of the results. AAF07478 to AAF11247  
CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents  
CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from  
CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from  
CC *Trichoderma reesei*, which are all specifically claimed in the present  
CC invention  
XX

XX Sequence 687 BP; 167 A; 190 C; 152 G; 178 T; 0 U; 0 Other;

XX Query Match 83.3%; Score 15; DB 3; Length 687;

XX Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Db 4 CACGCGACACACTTC 18  
531 CACGCGACACACTTC 545

XX RESULT 5

XX AAF1736/c ID AAF1736 standard; DNA; 747 BP.

XX AC AAF1736;

XX DT 30-APR-2001 (first entry)

XX DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:753.

XX Corynebacterium glutamicum; carbon metabolism and energy production;  
XX SMP protein; sugar metabolism and oxidative phosphorylation protein;  
XX fine chemical production; organic acid; proteinogenic amino acid;  
XX nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
XX nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
XX carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;  
XX diagnosis; Corynebacterium diptheriae; evolutionary study; ds.

XX OS Corynebacterium glutamicum.

XX PN WO200100844-A2.

XX PD 04-JAN-2001.

XX PF 23-JUN-2000; 2000WO-IB000943.

XX PR 25-JUN-1999; 99US-0141031P.

XX PI 08-JUN-1999; 99DE-01031412.

PR 08-JUL-1999; 99DE-01031413.  
PR 08-JUL-1999; 99DE-01031419.  
PR 08-JUL-1999; 99DE-01031420.  
PR 08-JUL-1999; 99DE-01031424.  
PR 08-JUL-1999; 99DE-01031428.  
PR 08-JUL-1999; 99DE-01031431.  
PR 08-JUL-1999; 99DE-01031433.  
PR 08-JUL-1999; 99DE-01031434.  
PR 08-JUL-1999; 99DE-01031510.  
PR 08-JUL-1999; 99DE-01031562.  
PR 08-JUL-1999; 99DE-01031534.  
PR 09-JUL-1999; 99DE-01032180.  
PR 09-JUL-1999; 99DE-01032227.  
PR 09-JUL-1999; 99DE-01032230.  
PR 09-JUL-1999; 99DE-0103208P.  
PR 14-JUL-1999; 99DE-01032924.  
PR 14-JUL-1999; 99DE-01032973.  
PR 14-JUL-1999; 99DE-01033005.  
PR 27-AUG-1999; 99DE-01040765.  
PR 31-AUG-1999; 99DE-0151572P.  
PR 03-SEP-1999; 99DE-01042076.  
PR 03-SEP-1999; 99DE-01042079.  
PR 03-SEP-1999; 99DE-01042086.  
PR 03-SEP-1999; 99DE-01042087.  
PR 03-SEP-1999; 99DE-01042088.  
PR 03-SEP-1999; 99DE-01042095.  
PR 03-SEP-1999; 99DE-01042123.  
PR 03-SEP-1999; 99DE-01042125.  
XX (BADT ) BASF AG.  
XX  
XX  
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;  
XX WPI; 2001-061975/07.  
XX P-PSDB; AAB79619.  
XX  
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar  
PT metabolism and oxidative phosphorylation protein for production or  
PT modulation of production of fine chemicals e.g. amino acids,  
PT carbohydrates or enzymes.  
XX  
XX  
PS Claim 3; Page 1209-1210; 1246pp; English.  
XX  
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar  
CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243  
CC to AAB 79633 which are involved in carbon metabolism and energy  
CC production. The C. glutamicum SMP gene can be used in vectors (II) for  
CC expression in host cells and production or modulation of production of  
CC fine chemicals, such as, an organic acid, a proteihogenic or  
CC nonproteihogenic amino acid (preferred), a purine or pyrimidine base, a  
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,  
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a  
CC polyetide, or an enzyme. The presence of (I) or SMP proteins (III)  
CC encoded by them are used for diagnosing the presence or activity of  
CC Corynebacterium diphteriae in a subject. (I), (II), (III) or host cells  
CC containing them are used to map genomes of organisms related to C.  
CC glutamicum, identify and localise C. glutamicum sequences of interest, in  
CC evolutionary studies, in determining SMP protein regions required for  
CC function, in modulating SMP protein activity, in modulating the  
CC metabolism of sugars, and in modulating high-energy molecule production  
CC in a cell (i.e. ATP, NADPH)  
XX  
SQ Sequence 747 BP; 175 A; 169 C; 205 G; 198 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 4; Length 747;  
Best Local Similarity 100.0%; Pred. No. 12;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CACCGACACACTTC 18  
|||  
DB 21 CACCGACACACTTC 7

RESULT 6  
AAH68534/C  
ID AAH68534 standard; DNA; 309400 BP.  
XX  
XX  
AC AAH68534;  
XX  
DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7069.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
KM organic acid synthesis; de.

OS Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX WPI; 2001-376931/40.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analyzing  
PT expression profile or pattern of a gene and identifying homologous gene.

XX Disclosure; SEQ ID NO 7069; 246pp + Sequence Listing; English.

XX  
XX  
PS The present invention provides a number of nucleotide and protein  
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
XX are useful for identifying the mutation point of a gene derived from a  
XX mutant of coryneform bacterium, measuring expression amount and analysing  
XX the expression profile or expression pattern of a gene derived from  
XX Coryneform bacterium, and identifying a homologue of a gene derived from  
XX coryneform bacterium. Coryneform bacteria are useful for producing amino  
XX acids, nucleic acids, vitamins, saccharides and organic acids,  
XX particularly L-lysine. The present sequence is a nucleic acid described  
XX in the exemplification of the invention. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from the European Patent Office

SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 5; Length 309400;  
Best Local Similarity 100.0%; Pred. No. 11;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CACCGACACACTTC 18  
|||  
DB 28188 CACCGACACACTTC 28174

RESULT 7

AAH68533/C  
ID AAH68533 standard; DNA; 349980 BP.

XX AAH68533;  
XX

DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7068.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

organic acid synthesis; ds.  
 Corynebacterium glutamicum.  
 EP1108790-A2.  
 20-JUN-2001.  
 18-DEC-2000; 2000EP-00127688.  
 16-DEC-1999; 99JP-00377484.  
 07-APR-2000; 2000JP-00159162.  
 03-AUG-2000; 2000JP-00280988.  
 (KYOW ) KYOWA HAKKO KOGYO KK.  
 Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 WPI; 2001-376931/40.  
 Novel polynucleotides derived from Coryneform bacteria, for identifying  
 mutation point of a gene, measuring expression of a gene, analyzing  
 expression profile or pattern of a gene and identifying homologous gene.  
 Disclosure; SEQ ID NO 7068; 246bp + Sequence Listing; English.  
 The present invention provides a number of nucleotide and protein  
 sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 are useful for identifying the mutation point of a gene derived from a  
 mutant of coryneform bacterium, measuring expression amount and analysing  
 the expression profile or expression pattern of a gene derived from  
 Coryneform bacterium, and identifying a homologue of a gene derived from  
 Coryneform bacterium. Coryneform bacteria are useful for producing amino  
 acids, nucleic acids, vitamins, saccharides and organic acids,  
 particularly L-lysine. The present sequence is a nucleic acid described  
 in the exemplification of the invention. Note: The sequence data for this  
 patent did not form part of the printed specification, but was obtained  
 in electronic format directly from the European Patent Office  
 Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;  
 Query Match 83.3%; Score 15; DB 5; Length 349980;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 4 CACGCGACACACTTC 18  
 328188 CACGCGACACACTTC 328174  
 RESULT 8  
 ADQ54932/C  
 ID ADQ54932 standard; DNA; 320 BP.  
 ADQ54932;  
 21-OCT-2004 (first entry)  
 Novel canine microarray-related DNA sequence SeqID6234.  
 canine microarray; drug screening; toxicity assay;  
 environmental pollutant; cellular response; gene expression profile;  
 toxic response; liver necrosis; fatty liver disease;  
 protein adduct formation; hepatitis; dog; ds.  
 Canis familiaris.  
 WO2004063324-A2.  
 29-JUL-2004.  
 05-MAY-2003; 2003WO-US013853.

03-MAY-2002; 2002US-0377240P.  
 (GENE-) GENE LOGIC INC.  
 (PFIZ ) PFIZER PROD INC.  
 Digiana JC, Porter M, Wei T;  
 WPI; 2004-561890/54.  
 New isolated nucleic acid molecule, useful for drug screening and  
 toxicity assays or for assessing the impact, including toxicity, of a  
 compound, pharmaceutical agent or environmental pollutant on a cell or  
 living organism.  
 Claim 1; SEQ ID NO 6234; 41bp; English.  
 This invention is related to a novel isolated canine nucleic acid  
 sequences and the construction of canine microarrays containing a  
 significant portion of the canine genome. The isolated canine nucleic  
 acid sequences of the invention may be useful for drug screening and  
 toxicity assays. The invention is therefore useful for assessing the  
 impact, including toxicity, of a compound, pharmaceutical agent or  
 environmental pollutant on a cell or living organism. The methods are  
 useful for detecting genes that are up- or down-regulated in canines in a  
 disease state. The sequences are useful as diagnostic agents or markers  
 to detect a cellular response in a sample individually or as part of a  
 gene expression profile. It is also useful as a target for agents that  
 modulate gene expression or activity. The database is useful for  
 producing electronic Northern blots that allow the user to determine the cell  
 type or tissue in which a given gene is expressed and to allow  
 determination of the abundance or expression level of a given gene in a  
 particular tissue or cell. The methods are useful for determining the  
 similarity of a toxic response to one or more individual compounds. The  
 methods are useful for predicting at least one toxic response or the  
 likelihood that a compound or test agent will induce various specific  
 pathologies such as those of the liver (liver necrosis, fatty liver  
 disease, protein adduct formation or hepatitis), those of the kidney,  
 heart, brain or testes, or other pathologies associated with at least one  
 of the toxins. The methods are also useful for predicting or elucidating  
 the potential cellular pathways influenced, induced or modulated by the  
 compound or test agent due to the similarity of the expression profile  
 compared to the profile induced by a known toxin. The present sequence is  
 that of a canine DNA sequence which was claimed for use during the  
 production of a canine microarray of the invention.  
 Sequence 320 BP; 35 A; 97 C; 134 G; 49 T; 0 U; 5 Other;  
 Query Match 77.8%; Score 14; DB 13; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 ACCACGCGACACAC 15  
 230 ACCACGCGACACAC 217  
 RESULT 9  
 ABV40711/C  
 ID ABV40711 standard; cDNA; 435 BP.  
 ABV40711;  
 16-SEP-2002 (first entry)  
 Human prostate expression marker cDNA 40702.  
 Human, prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 pharmacogenomic marker; gene; ss.  
 Homo sapiens.  
 WO200160860-A2.

XX 23-AUG-2001.  
PD 20-FEB-2001; 2001WO-US005171.  
XX  
PF 17-FEB-2000; 2000US-0183319P.  
XX PR 16-MAR-2000; 2000US-0189862P.  
XX PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 8202; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 435 BP; 128 A; 102 C; 110 G; 95 T; 0 U; 0 Other;  
XX  
Query Match 77.8%; Score 14; DB 5; Length 435;  
Best Local Similarity 100.0%; Pred. NO. 50;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CACGCGACACACTT 17  
DB 286 CACGCGACACACTT 273  
XX  
RESULT 10  
ABV31743/c  
ID ABV31743 standard; cDNA; 435 BP.  
XX  
AC ABV31743;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 31734.  
XX  
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KM pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.

PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 6811; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 435 BP; 128 A; 102 C; 110 G; 95 T; 0 U; 0 Other;  
XX  
Query Match 77.8%; Score 14; DB 5; Length 435;  
Best Local Similarity 100.0%; Pred. NO. 50;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CACGCGACACACTT 17  
DB 286 CACGCGACACACTT 273  
XX  
RESULT 11  
ABN75645  
ID ABN75645 standard; cDNA; 435 BP.  
XX  
AC ABN75645;  
XX  
DT 08-JUL-2002 (first entry)  
XX  
DE Human ORF592 cDNA, SEQ ID NO:1183.  
XX  
XX Human; ORF, open reading frame; ORF; drug screening; diagnosis;  
KM disease monitoring; cytokine; cell proliferation; cell differentiation;  
KM immune modulation; haematopoiesis regulation; tissue growth;  
KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;  
KM thrombolytic; tumour inhibition; bodily characteristics; fertility;  
KM behaviour; cancer; proliferative disorder; neurological disorder;  
KM cardiovascular disease; immune system disorder; organ transplantation;  
KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
KM hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
KM vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;  
KM neuroprotective; antineurosecretory; anticegulatory; thrombolytic;  
KM cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
KM dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200190366-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 24-MAY-2001; 2001WO-US017076.  
XX  
PR 24-MAY-2000; 2000US-0206690P.  
XX





PT from chemically treated DNA.

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ13410-  
 CC ABO54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention

XX Sequence 732 BP; 245 A; 341 C; 82 G; 64 T; 0 U; 0 Other;

XX Query Match 77.8%; Score 14; DB 6; Length 732;

PS Best Local Similarity 100.0%; Pred. No. 49;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACCACGGGACAC 15

Db 185 ACCACGGGACAC 198

Search completed: February 10, 2005, 13:44:08

Job time : 22.6055 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41, Search time 29.935 Seconds  
(without alignments)  
3460.943 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18

Sequence: 1 tgcgtcttcgtcgtacttc 18

Scoring table: OLIGO NUC  
Gapop 60.0, Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 18    | 100.0       | 18     | 9     | US-09-872-051-11     |
| 2          | 18    | 100.0       | 18     | 18    | US-10-790-430-11     |
| 3          | 18    | 100.0       | 1183   | 9     | US-09-872-051-8      |
| 4          | 18    | 100.0       | 1183   | 18    | US-10-790-430-8      |
| 5          | 18    | 88.9        | 1818   | 18    | US-10-425-115-84576  |
| 6          | 15    | 83.3        | 21     | 17    | US-10-349-143-7605   |
| 7          | 15    | 83.3        | 445    | 18    | US-10-357-930-19880  |
| 8          | 15    | 83.3        | 500    | 18    | US-10-357-930-49671  |
| 9          | 15    | 83.3        | 533    | 17    | US-10-027-632-234880 |
| 10         | 15    | 83.3        | 533    | 17    | US-10-027-632-234880 |
| 11         | 15    | 83.3        | 641    | 18    | US-10-767-701-6161   |

|    |      |       |    |                      |                    |
|----|------|-------|----|----------------------|--------------------|
| 12 | 83.3 | 695   | 17 | US-10-424-599-77613  | Sequence 77613, A  |
| 15 | 83.3 | 3167  | 10 | US-09-764-891-8948   | Sequence 8948, Ap  |
| 13 | 83.3 | 70019 | 18 | US-10-323-281-823    | Sequence 823, App  |
| 14 | 77.8 | 20    | 17 | US-10-289-782-5457   | Sequence 5457, Ap  |
| 15 | 77.8 | 25    | 18 | US-10-717-597-3196   | Sequence 3196, Ap  |
| 16 | 77.8 | 25    | 18 | US-10-717-597-3197   | Sequence 3197, Ap  |
| 17 | 77.8 | 149   | 18 | US-10-357-930-15466  | Sequence 15466, A  |
| 18 | 77.8 | 154   | 18 | US-10-357-930-15388  | Sequence 15388, A  |
| 19 | 77.8 | 169   | 9  | US-09-864-761-26676  | Sequence 26676, A  |
| 20 | 77.8 | 201   | 18 | US-10-741-601-5342   | Sequence 5342, Ap  |
| 21 | 77.8 | 201   | 18 | US-10-741-601-5350   | Sequence 5350, Ap  |
| 22 | 77.8 | 201   | 18 | US-10-741-601-5352   | Sequence 5352, Ap  |
| 23 | 77.8 | 201   | 18 | US-10-741-601-20494  | Sequence 20494, A  |
| 24 | 77.8 | 230   | 18 | US-10-357-930-36207  | Sequence 36207, A  |
| 25 | 77.8 | 241   | 17 | US-10-305-720-403    | Sequence 403, App  |
| 26 | 77.8 | 257   | 18 | US-10-674-124A-21630 | Sequence 21630, A  |
| 27 | 77.8 | 366   | 17 | US-10-282-122A-31878 | Sequence 31878, A  |
| 28 | 77.8 | 379   | 18 | US-10-357-930-36285  | Sequence 36285, A  |
| 29 | 77.8 | 400   | 10 | US-09-918-995-33608  | Sequence 33608, A  |
| 30 | 77.8 | 400   | 17 | US-10-242-535A-20475 | Sequence 20475, A  |
| 31 | 77.8 | 400   | 17 | US-10-085-783A-20475 | Sequence 20475, A  |
| 32 | 77.8 | 426   | 17 | US-10-282-122A-31878 | Sequence 31878, A  |
| 33 | 77.8 | 435   | 17 | US-10-042-535A-38277 | Sequence 38277, A  |
| 34 | 77.8 | 435   | 17 | US-10-085-783A-38277 | Sequence 38277, A  |
| 35 | 77.8 | 441   | 17 | US-10-242-535A-25043 | Sequence 25043, A  |
| 36 | 77.8 | 441   | 17 | US-10-085-783A-25043 | Sequence 25043, A  |
| 37 | 77.8 | 450   | 17 | US-10-424-599-110092 | Sequence 110092, A |
| 38 | 77.8 | 483   | 16 | US-10-029-386-23957  | Sequence 23957, A  |
| 39 | 77.8 | 496   | 9  | US-09-864-761-10034  | Sequence 10034, A  |
| 40 | 77.8 | 496   | 18 | US-10-357-930-45260  | Sequence 45260, A  |
| 41 | 77.8 | 496   | 18 | US-10-357-930-45326  | Sequence 45326, A  |
| 42 | 77.8 | 515   | 11 | US-09-864-408A-8531  | Sequence 8531, Ap  |
| 43 | 77.8 | 547   | 16 | US-10-029-386-10240  | Sequence 10240, A  |
| 44 | 77.8 | 549   | 10 | US-09-991-935-126    | Sequence 126, App  |
| 45 | 77.8 | 573   | 18 | US-10-021-323-14061  | Sequence 14061, A  |

#### ALIGNMENTS

RESULT 1

US-09-872-051-11

Sequence 11, Application US/09872051

Patent No. US20020013960A1

GENERAL INFORMATION:

APPLICANT: Monsanto Co

APPLICANT: Behr, Carl

APPLICANT: Hironaka, Catherine

APPLICANT: Heck, Gregory

APPLICANT: You, Jinsong

TITLE OF INVENTION: Corn Event PV-ZMGR32 (n603) and Composition and Methods for Det

FILE REFERENCE: 38-21(52258)B

CURRENT APPLICATION NUMBER: US/09/872,051

PRIOR FILING DATE: 2001-06-01

PRIOR APPLICATION NUMBER: 60/213,567

PRIOR FILING DATE: 2000-06-22

PRIOR APPLICATION NUMBER: 60/241,215

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: 60/240,014

PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.0

SEQ ID NO 11

LENGTH: 18

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: source

LOCATION: (1) ..(18)

OTHER INFORMATION: Zea maize genomic DNA and vector DNA

US-09-872-051-11

Query Match 100.0%; Score 18; DB 9; Length 18;

Best Local Similarity 100.0%; Pred. No. 0.97;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 1 TGCTGTTCTGCTGACTTT 18

## RESULT 2

US-10-790-430-11  
; Sequence 11, Application US/10790430  
; Publication No. US20040139493A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/10/790,430  
; CURRENT FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 18  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(18)  
; OTHER INFORMATION: Zea maize genomic DNA and vector DNA  
US-10-790-430-11

Query Match 100.0%; Score 18; DB 18; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.97;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 1 TGCTGTTCTGCTGACTTT 18

## RESULT 3

US-09-872-051-8  
; Sequence 8, Application US/09872051  
; Patent No. US20020013960A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/09/872,051  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13

NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(1183)  
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
; OTHER INFORMATION: 165-381 construct vector DNA  
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpsA  
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA  
US-09-872-051-8

Query Match 100.0%; Score 18; DB 9; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.88;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 678 TGCTGTTCTGCTGACTTT 695

## RESULT 4

US-10-790-430-8  
; Sequence 8, Application US/10790430  
; Publication No. US20040139493A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/10/790,430  
; CURRENT FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(1183)  
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
; OTHER INFORMATION: 165-381 construct vector DNA  
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpsA  
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA  
US-10-790-430-8

Query Match 100.0%; Score 18; DB 18; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.88;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGTTCTGCTGACTTT 18  
|||||  
Db 678 TGCTGTTCTGCTGACTTT 695

## RESULT 5

US-10-425-115-84576

```
/ Sequence 84576, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 84576
/ LENGTH: 1818
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(1818)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_177144C.1
US-10-425-115-84576
```

```
Query Match      88.9%; Score 16; DB 18; Length 1818;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGCTGTTCTGCTGACT 16
          |||||
DB      59 TGCTGTTCTGCTGACT 74
```

```
RESULT 6
US-10-349-143-7605
/ Sequence 7605, Application US/10349143
/ Publication No. US20040005584A1
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marla
/ APPLICANT: Chumakov, Ilya
/ TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
/ FILE REFERENCE: GENSET.020CP1
/ CURRENT APPLICATION NUMBER: US/10/349,143
/ CURRENT FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US/09/422,978
/ PRIOR FILING DATE: 1999-10-20
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
/ PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
/ NUMBER OF SEQ ID NOS: 11796
/ SEQ ID NO 7605
/ LENGTH: 21
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: primer_bind
/ LOCATION: 1..21
/ OTHER INFORMATION: upstream amplification primer 99-9623 for SEQ 3671,
US-10-349-143-7605
```

```
Query Match      83.3%; Score 15; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CTGTTCTGCTGACTT 17
          |||||
DB      2 CTGTTCTGCTGACTT 16
```

```
RESULT 7
US-10-357-930-19880/c
/ Sequence 19880, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 19880
/ LENGTH: 445
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-357-930-19880
```

```
Query Match      83.3%; Score 15; DB 18; Length 445;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGCTGTTCTGCTGAC 15
          |||||
DB      256 TGCTGTTCTGCTGAC 242
```

```
RESULT 8
US-10-357-930-49671/c
/ Sequence 49671, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
```

;; PRIOR FILING DATE: 2000-12-13  
;; NUMBER OF SEQ ID NOS: 62232  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 49671  
;; LENGTH: 500  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-357-930-49671

Query Match 83.3%; Score 15; DB 18; Length 500;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGAC 15  
Db 296 TGCTGTTCTGCTGAC 282

RESULT 9  
US-10-027-632-234880/c  
; Sequence 234880, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 234880  
; LENGTH: 533  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-234880

Query Match 83.3%; Score 15; DB 13; Length 533;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TGTTCTGCTGACTT 18  
Db 374 TGTTCTGCTGACTT 360

RESULT 10  
US-10-027-632-234880/c  
; Sequence 234880, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12

;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 234880  
;; LENGTH: 533  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-234880

Query Match 83.3%; Score 15; DB 17; Length 533;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TGTTCTGCTGACTT 18  
Db 374 TGTTCTGCTGACTT 360

RESULT 11  
US-10-767-701-6161  
; Sequence 6161, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 6161  
; LENGTH: 641  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS35548\_1  
US-10-767-701-6161

Query Match 83.3%; Score 15; DB 16; Length 641;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTGTTCTGCTGAC 15  
Db 586 TGCTGTTCTGCTGAC 600

RESULT 12  
US-10-424-599-77613  
; Sequence 77613, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53523)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

```
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 77613
/ LENGTH: 695
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_41100C.1
US-10-424-599-77613

Query Match      83.3%; Score 15; DB 17; Length 695;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGCTGTTCTGCTGAC 15
         |||||
Db      187 TGCTGTTCTGCTGAC 201

RESULT 13
US-09-764-891-8948
/ Sequence 8948, Application US/09764891
/ Publication No. US20030077808A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC006
/ CURRENT APPLICATION NUMBER: US/09/764,891
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 10231
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 8948
/ LENGTH: 3167
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-764-891-8948

Query Match      83.3%; Score 15; DB 10; Length 3167;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 TGTTCTGCTGACTTT 18
         |||||
Db      1739 TGTTCTGCTGACTTT 1753

RESULT 14
US-10-322-281-823
/ Sequence 823, Application US/10322281
/ Publication No. US20040126762A1
/ GENERAL INFORMATION:
/ APPLICANT: David W. Morris
/ APPLICANT: Marc S. Malandro
/ TITLE OF INVENTION: Novel Compositions and Methods in Cancer
/ FILE REFERENCE: 529452001000
/ CURRENT APPLICATION NUMBER: US/10/322,281
/ CURRENT FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 823
/ LENGTH: 70019
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-10-322-281-823

Query Match      83.3%; Score 15; DB 18; Length 70019;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 TGTTCTGCTGACTTT 18
         |||||
Db      27137 TGTTCTGCTGACTTT 27151
```

```
RESULT 15
US-10-289-762-5457/c
/ Sequence 5457, Application US/10289762
/ Publication No. US20040006218A1
/ GENERAL INFORMATION:
/ APPLICANT: Grifais, R
/ TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragmen-
/ thereof and uses thereof, in particular for the diagnosis, pre-
/ FILE REFERENCE: 9710-003-999
/ CURRENT APPLICATION NUMBER: US/10/289,762
/ CURRENT FILING DATE: 2003-03-27
/ NUMBER OF SEQ ID NOS: 6849
/ SEQ ID NO 5457
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Chlamydia pneumoniae
US-10-289-762-5457

Query Match      77.8%; Score 14; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 GTTCTGCTGACTTT 18
         |||||
Db      16 GTTCTGCTGACTTT 3
```

Search completed: February 9, 2005, 11:12:46  
Job time : 30.935 secs

**This Page Blank (uspto)**



ORGANISM      synthetic construct  
 other sequences; artificial sequences.

REFERENCE  
 1 Behr, C.F., Hironaka, C., Heck, G.R. and You, J.  
 Corn transformant pv-zmgt32 (hk603) and compositions and methods  
 for detection thereof  
 Patent: EP 1167531-A 8 02-JAN-2002;  
 Monsanto Technology LLC (US)

JOURNAL  
 location/Qualifiers

FEATURES  
 source  
 1. 1183  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="1-164 Agrobacterium tumefaciens nos 3' terminator  
 165-381 construct vector DNA 382-686 Zea mays genomic DNA"  
 genes, rpm1 and rpm2 687-1183 Zea mays genomic DNA"

ORIGIN

Query Match      100.0%; Score 18; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 9.9;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  
 1 ACCAAGCTTTATATAG 18  
 |||||  
 156 ACCAAGCTTTATATAG 173

RESULT 3  
 AC106676      225438 bp      DNA      linear      HTG 13-MAY-2003  
 LOCUS      Rattus norvegicus clone CH230-107N12, WORKING DRAFT SEQUENCE, 3  
 DEFINITION      unoriented pieces.  
 AC106676  
 AC106676 4 GI:30578541  
 HTG; HTGS\_PHAASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 225438)  
 Muzny, D., Maric, M., Mettler, M., Lee, A., Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,  
 Ayalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C. M., Gabisl, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gubregovis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guayra, W.,  
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensuawe, L., Louisedge, H., Lozada, R. J., Lu, X., Ma, J.,  
 Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangun, A.,  
 Mahoney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
 Mloosebajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Natir, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwaokemele, O., Okwuonu, G., Olarinmase, A., Pal, S., Parke, K.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 Project Information  
 Center project name: GLUT  
 Center clone name: CH230-107N12  
 Summary Statistics  
 Assembly program: Atlas 3.0;  
 Consensus quality: 201968 bases at least Q40  
 Consensus quality: 205320 bases at least Q30  
 Consensus quality: 207660 bases at least Q20  
 Estimated insert size: 213439; sum-of-coverage estimation  
 Quality coverage: 7x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 NOTE: This is a "working draft" sequence. It currently  
 consists of 3 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 222660: contig of 222660 bp in length  
 \* 222661 222760: gap of unknown length  
 \* 222761 224015: contig of 1255 bp in length



```

FEATURES
  source
    * 224016 224115: gap of unknown length
    * 224116 225438: contig of 1323 bp in length
    Location/Qualifiers
      1..225438
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-107N12"
        1..1186
          /note="wgs_end_extension
          clone_end:5p6"
          4773..5453
            /note="clone boundary
            clone_end:5p6
            site:ECORI
            end_sequence:BH2711428"
            complement(121111..221919)
              /note="clone boundary
              clone_end:T7
              site:ECORI
              end_sequence:BH2711425"

ORIGIN

Query Match      100.0%; Score 18; DB 2; Length 225438;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 ACCAAGCTTTTATATAG 18
        |||||||
Db      13992 ACCAAGCTTTTATATAG 14009

RESULT 4
LOCUS      AC110467
DEFINITION Rattus norvegicus clone CH230-320D17, *** SEQUENCING IN PROGRESS
VERSION     AC110467.4 GI:23195262
AUTHORS     HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS    Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 232994)
            Murzu D.Marie, Metzger M.Lee, Abramson S., Adams C., Alder J.,
            Allen C., Allen H., Alb Brooks S., Amin A., Angliano D.,
            Anyalebech V., Aoyagi A., Ayodeji W., Baca E., Baden H.,
            Baldwin D., Bandaramaiki D., Barber M., Barnstead M.,
            Biawalo K., Blair J., Blankenburg K., Blyth P., Brown M.,
            Bryant N., Buhay C., Burch P., Butrell K., Calderon E.,
            Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
            Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
            Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., P.Souza L.,
            Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
            Delgado O., Denison S., Deramo C., Ding Y., Dinh H., Divya K.,
            Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
            Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
            Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
            Frazer C.M., Gabisi A., Ganter R., Garcia A., Garner T., Garza M.,
            Gebregorjsgis E., Geer K., Gilli R., Grady M., Guerra W., Guevara W.,
            Gunaratne P., Haaland W., Hamill C., Hamilton C., Hamilton K.,
            Harvey Y., Haylak P., Hawes A., Henderson N., Hernandez J.,
            Hernandez R., Hines S., Hulyk S., Hune J., Idlebird D., Jackson A.,
            Hollins B., Howells S., Hulyk S., Hune J., Idlebird D., Jackson A.,
            Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolyvec A.,
            Karachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
            Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
            Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
            Lounsbury L., Louised H., Lozao R.J., Lu X., Ma J.,
            Maheshwari M., Mahindaratne M., Mahmoud M., Malloy K., Mangum A.,
            Mangum B., Mapa P., Martin K., Martin R., Martine E.,

```

Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munida, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,  
Pnackelmeier, O., Okunolu, G., Olarinmoye, A., Pal, S.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Planchet, C.,  
Plummer, P., Polidexter, A., Popovic, D., Primm, E., Pu, L.,  
Pizzo, M., Quirio, J., Rachlin, E., Reeves, K., Reigh, M.A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,  
Sneid, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,  
Valas, R., Varga, V., Villaseña, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willison, R., Wlezyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Neudorfer, G., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G., and Gibbs, R.A.  
  
Unpublished  
Direct Submission  
2 (bases 1 to 232994)  
  
Worley, K.C.  
Direct Submission  
Submitted (13-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 232994)  
  
Rat Genome Sequencing Consortium.  
  
Direct Submission  
Submitted (19-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Sep 19, 2002 this sequence version replaced gi:21741421.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). As a result, the  
sequence may extend beyond the ends of the clone and there may be  
contigs that consist entirely of whole genome shotgun sequence  
reads. Both end sequences and whole genome shotgun sequence only  
contigs will be indicated in the feature table.  
  
----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
  
----- Project Information -----  
Center project name: GRNM  
Center clone name: CH230-320D17  
----- Summary Statistics -----  
Assembly program: Phrap; version 0.9903129  
Consensus quality: 217650 bases at least Q40  
Consensus quality: 220144 bases at least Q30  
Consensus quality: 221530 bases at least Q20  
Estimated insert size: 235091; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation  
  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html))  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
  
1 152446: contig of 152446 bp in length  
\* 152447 152546: gap of unknown length  
\* 152547 214193: contig of 61647 bp in length

```

* 214194 214293: gap of unknown length
* 214294 215380: contig of 1087 bp in length
* 215381 215480: gap of unknown length
* 215481 217850: contig of 2370 bp in length
* 217851 217950: gap of unknown length
* 217951 218965: contig of 1015 bp in length
* 218966 219065: gap of unknown length
* 219066 220974: contig of 1909 bp in length
* 220975 221075: gap of unknown length
* 221075 223294: contig of 11920 bp in length.
Location/Qualifiers
1. .232994
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-320D17"
5849..7480
/note="wgs_contig"

FEATURES
source
misc_feature

ORIGIN
Query Match 100.0%; Score 18; DB 2; Length 232994;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATATAG 18
|||||
Db 83203 ACCAAGCTTTATATAG 83220

RESULT 5
AC138014 118174 bp DNA linear HTG 04-SEP-2004
LOCUS Medicago truncatula clone mth2-17p13, WORKING DRAFT SEQUENCE, 3
DEFINITION Medicago truncatula (barrel medic)
AC138014
AC138014.30 GI:51689812
HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Medicago truncatula
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 118174)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B. A.
Medicago truncatula BAC Clone mth2-17p13
Unpublished
2 (bases 1 to 118174)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B. A.
Direct Submission
Submitted (10-DEC-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 118174)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B. A.
Direct Submission
Submitted (04-SEP-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Sep 4, 2004 this sequence version replaced gi:51699588.
Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

```

```

* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 78995: contig of 78995 bp in length
78996 79095: gap of unknown length
79096 84864: contig of 5769 bp in length
84865 84964: gap of unknown length
84965 118174: contig of 33210 bp in length.
Location/Qualifiers
1. 118174
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-17p13"
/clone_11b="Medicago truncatula BAC library H2"

FEATURES
source

ORIGIN
Query Match 94.4%; Score 17; DB 2; Length 118174;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATATATA 17
|||||
Db 70329 ACCAAGCTTTATATATA 70345

RESULT 6
AC004925 154959 bp DNA linear PRI 30-JUN-2004
LOCUS Homo sapiens PAC clone RP5-907C10 from 7, complete sequence.
AC004925
AC004925.1 GI:4156174
HTG.
Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 154959)
Hillier, L.W., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H.,
Magnez-McPherson, C., Layman, D., Maas, J., Jaeger, S., Walker, R.,
Wylie, K., Sekhon, M., Becker, M.C., O'Laughlin, M.D., Schaller, M.E.,
Fewell, G.A., Delehaney, K.D., Miner, T.L., Nash, W.E., Cordes, M.,
Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., Ali, J., Andrews, S.,
Isak, A., Vanbrunt, A., Nguyen, C., Du, F., Lamar, B., Courtney, L.,
Kalkbrenner, J., Ozersky, P., Bielicki, L., Scott, K., Holmes, A.,
Harkins, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C.,
Dauphin-Kohlberg, S., Kozlowicz-Reilly, A., Leonard, S., Kohlfing, T.,
Rock, S.M., Tin-Wollam, A.M., Abbott, A., Minx, P., Maupin, R.,
Stromwater, C., Latreille, P., Miller, N., Johnson, D., Murray, J.,
Moessner, J.P., Wendt, M.C., Yang, S.P., Schultz, B.R., Mallis, J.W.,
Spieth, J., Bieri, T.A., Nelson, J.O., Berkowicz, N., Wohlmann, P.E.,
Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Bedell, J.A.,
Mardis, E.R., Clifton, S.W., Chissole, S.L., Marra, W.A., Raymond, C.,
Haugen, E., Gillett, W., Zhou, Y., James, R., Phelps, K., Iadonato, S.,
Bubb, K., Simms, E., Levy, R., Clendenning, J., Kaul, R., Kent, W.D.,
Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrente, D.,
Chinwalla, A.T., Gish, W.R., Bddy, S.R., McPherson, J.D., Olson, M.V.,
Eichler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.
The DNA sequence of human chromosome 7
Nature 424 (6945), 157-164 (2003)
22737999
MEDLINE
12853948
2 (bases 1 to 154959)
Jones, K., Keppeler, D. and Tin-Wollam, A.
The sequence of Homo sapiens PAC clone RP5-907C10
Unpublished (2001)
3 (bases 1 to 154959)
Waterston, R.H.
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MO 63108, USA  
4 (bases 1 to 154959)  
Waterson,R.H.  
Direct Submission  
Submitted (14-JAN-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 154959)  
Waterson,R.  
Direct Submission  
Submitted (22-JAN-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
6 (bases 1 to 154959)  
Waterson,R.  
Direct Submission  
Submitted (21-DEC-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
7 (bases 1 to 154959)  
Waterson,R.  
Direct Submission  
Submitted (26-APR-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
8 (bases 1 to 154959)  
Wilson,R.  
Direct Submission  
Submitted (30-JAN-2004) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jan 14, 1999 this sequence version replaced gi:3213072.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: saplens@watson.wustl.edu  
----- Summary Statistics  
Center project name: H\_DJ0907C10

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and  
sequencing collaboration between the NHGRI Chromosome 7 Mapping  
Project (Eric D. Green, Director), John D. McPherson in the  
Department of Genetics (Washington University), and the Washington  
University Genome Sequencing Center. For additional information  
about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send  
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

SOURCE INFORMATION:  
This clone was derived from human PAC library RPCI-5, prepared by  
Pieter de Jong and coworkers at http://www.chori.org using the  
method described by Ioannou et al., Nature Genetics 6:84-9 (1994).  
The library is from one male donor.  
The clone may be obtained either from Genome Systems, Inc.  
(http://www.genomesystems.com) or Research Genetics, Inc.  
(http://www.resgen.com); or from Pieter de Jong.  
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:  
Actual start of this clone is at base position 1 of RPS-907C10  
actual end is at 154959 of RPS-907C10.

| FEATURES      | Location/Qualifiers   |
|---------------|---|
| source        | 1..154959<br>/organism="Homo sapiens"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:9606"<br>/chromosome="7"<br>/map="7"  |
| repeat_region | /clone="RPS-907C10"<br>/clone_1ib="RPCI-5"<br>68..391<br>/rpt_family="L1"   |
| repeat_region | 731..752<br>/rpt_family="AT-rich"   |
| repeat_region | 1128..1429<br>/rpt_family="Alu"   |
| repeat_region | 1554..1648<br>/rpt_family="L1"  |
| repeat_region | 1659..1786<br>/rpt_family="L1"  |
| repeat_region | 1790..1936<br>/rpt_family="L1"  |
| repeat_region | 1949..2296<br>/rpt_family="L1"  |
| repeat_region | 2297..2595<br>/rpt_family="Alu"   |
| repeat_region | 2596..2760<br>/rpt_family="L1"  |
| gene          | 2814..26553<br>/gene="PORT1"  |
| mRNA          | join(2814..2870,6849..7005,8634..8839,14398..14533,<br>20470..20556,22507..22598,24455..24560,25738..26553)<br>/gene="PORT1"  |
| mRNA          | join(2814..2870,6849..7005,8634..8839,14398..14533,<br>20470..20556,22507..22598,24455..24560,25738..26553)<br>/gene="PORT1"  |
| mRNA          | join(2814..2870,6849..7005,8634..8839,14398..14533,<br>20470..20556,22507..22598,24455..24560,25738..26553)<br>/gene="PORT1"  |
| CDS           | join(2814..2870,6849..7005,8634..8839,14398..14533,<br>20470..20556,22507..22598,24455..24560,25738..26553)<br>/codon_start=3<br>/product="unknown"<br>/protein_id="AAD08852.1"<br>/db_xref="GI:4176376"<br>/translation="SGSVSLYEVCQQLSATLTTHQVLEFPLCAILKQAPQIR<br>IRAKLRKRPRLRFQSVKLCPCPKHLOEVPBHGDDLIIFODGATKTPDVGLQVTSLY<br>DSKLTWKNGKGRKRVAVFVKNNGILPLSNCELLIEGTLSEICKLSNKNSVYIPVR<br>SGHDELFLDLSAPFLIGTIIHHYGCCKSLSTSIONLSLYDKTSMIPSSVAAAGI<br>VPLQYVFMFTLDDGTGVLBAFLMDSKRFQIPASVLMDDDLQSKVDIMDKFCPP<br>GIKIDAVFMCECFKSYNTNGTNDQICVOIFPTTVADVI" |
| repeat_region | 3436..4963<br>/rpt_family="L1"  |
| repeat_region | 5057..5355<br>/rpt_family="Alu"   |
| repeat_region | 5384..5685<br>/rpt_family="Alu"   |
| repeat_region | 6265..6308<br>/rpt_family="CT-rich"   |
| repeat_region | 6628..6748<br>/rpt_family="L2"  |
| repeat_region | 7101..7121<br>/rpt_family="AT-rich"   |
| repeat_region | 7421..7446<br>/rpt_family="AT-rich"   |
| repeat_region | 8020..8148<br>/rpt_family="MIR"   |
| repeat_region | 8174..8194<br>/rpt_family="AT-rich"   |
| repeat_region | 9616..9810<br>/rpt_family="L1"  |
| repeat_region | 9834..9908  |

```
repeat_region /rpt_family="L1"
9941..10129 /rpt_family="L1"
repeat_region 10420..10528
repeat_region 10528..10907 /rpt_family="L1"
repeat_region 10908..11201 /rpt_family="L1"
repeat_region 11202..12388 /rpt_family="L1u"

Query Match 94.4%; Score 17; DB 9; Length 154959;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTATATAG 18
DB 30174 CCAAGCTTTATATAG 30190
```

```
RESULT 7
AL133551 175940 bp DNA linear PRI 21-AUG-2000
LOCUS Human DNA sequence from clone RP11-57G10 on chromosome 10 contains
a J-domain containing protein (JDBP) isoform B, the SIR1 gene
(Sir2-like proteins (siruitins) type 1), part of a novel gene
similar to KIAA0032, two ribosomal pseudogenes, 2 Cpg islands,
ESTs, STSs and GSSs, complete sequence.
AL133551
ACCESSION AL133551.13 GI:9407715
VERSION HTG; J-domain; RPL21; SIR1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175940)
Wilson, S.
Direct Submission
Submitted (21-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
Requester: clonerequest@sanger.ac.uk
On Jul 23, 2000 this sequence version replaced gi:9367364.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unure'
feature key.
```

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
Em: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr10>  
RP11-57G10 is from the library RPCR-11.1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>  
VECTOR: pBACe3.6  
This sequence is the entire insert of clone RP11-57G10 The true left end of clone RP11-474D14 is at 172652 in this sequence.

## FEATURES

```
source 1..175940
location/Qualifiers
1..175940
/organism="Homo sapiens"
/mol_type="genomic DNA"
```

```
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-57G10"
/clone_1kb="RPCR-11.1"
1..2318
/note="L1P10 repeat: matches 3815..6158 of consensus"
2338..2767
/note="L1M4 repeat: matches 4946..5385 of consensus"
2768..3058
/note="AluDo repeat: matches 1..296 of consensus"
3059..3371
/note="L1M4 repeat: matches 5385..5704 of consensus"
3372..3682
/note="AluSx repeat: matches 1..310 of consensus"
3683..4322
/note="L1M4 repeat: matches 5704..6238 of consensus"
4346..4707
/note="L1M5 repeat: matches 5798..6175 of consensus"
4872..4926
/note="MIR repeat: matches 204..258 of consensus"
4880..4938
/note="L2 repeat: matches 2648..2731 of consensus"
5029..5146
/note="MIR1 repeat: matches 436..564 of consensus"
5150..5194
/note="MIR2B repeat: matches 2..50 of consensus"
5195..5754
/note="MIR2A repeat: matches 1..453 of consensus"
5755..6125
/note="MIR2B repeat: matches 42..454 of consensus"
6142..6207
/note="33 copies 2 mer ac 86% conserved"
7403..7494
/note="MIR1 repeat: matches 1..81 of consensus"
7561..7932
/note="MIR1 repeat: matches 180..568 of consensus"
8283..8422
/note="MIR7A repeat: matches 191..336 of consensus"
8415..8907
/note="MIR1A repeat: matches 11..527 of consensus"
8908..9107
/note="MIR7A repeat: matches 1..211 of consensus"
9110..9171
/note="MIR repeat: matches 76..146 of consensus"
9449..9750
/note="AluY repeat: matches 1..304 of consensus"
10489..10938
/note="LIR2 repeat: matches 1..448 of consensus"
11938..12442
/note="Cpg island"
/evidence="not experimental"
12362..12436
/note="TCA repeat: matches 8..82 of consensus"
12947..13250
/note="AluY repeat: matches 1..303 of consensus"
13888..13947
/note="MIR1 repeat: matches 453..514 of consensus"
14085..14387
/note="AluSd repeat: matches 1..304 of consensus"
complement(14553..14948)
/note="match: GSS: Em:Q0113044"
14626..14764
/note="MIR1 repeat: matches 229..366 of consensus"
15268..15328
/note="MIR1 repeat: matches 113..172 of consensus"
15822..16120
/note="AluSx repeat: matches 1..306 of consensus"
16150..16463
/note="AluDb repeat: matches 1..312 of consensus"
17598..18107
/note="L2 repeat: matches 1843..2330 of consensus"
18101..18225
/note="match: GSS: Em:AQ627100"
```

```

misc_feature      18230..18711      /note="match: GSS: Em:AQ6727100"
repeat_region     19538..19830      /note="AluX repeat: matches 1..294 of consensus"
repeat_region     20049..20256      /note="MERSA repeat: matches 12..224 of consensus"
repeat_region     20724..20880      /note="MIR repeat: matches 20..175 of consensus"
repeat_region     21560..21844      /note="AluX repeat: matches 1..302 of consensus"
repeat_region     21917..22003      /note="L2 repeat: matches 2207..2295 of consensus"
misc_feature      21929..22382      /note="match: GSS: Em:AQ672646"
repeat_region     22135..22178      /note="L2 repeat: matches 1..305 of consensus"
repeat_region     22288..22595      /note="L2 repeat: matches 1..305 of consensus"
repeat_region     24457..24579      /note="L2 repeat: matches 1..305 of consensus"
repeat_region     24729..24815      /note="L2 repeat: matches 125..213 of consensus"
repeat_region     24992..25152      /note="L2 repeat: matches 10..120 of consensus"
repeat_region     25153..25419      /note="L2 repeat: matches 1..263 of consensus"
repeat_region     25420..25432      /note="L2 repeat: matches 120..133 of consensus"
repeat_region     26638..26802      /note="MERSA repeat: matches 7..189 of consensus"
repeat_region     27004..27072      /note="MIR repeat: matches 238..306 of consensus"
repeat_region     27063..27175      /note="MIR repeat: matches 144..250 of consensus"
repeat_region     27176..27447      /note="L2 repeat: matches 3..274 of consensus"
repeat_region     27453..27532      /note="L2 repeat: matches 2586..2705 of consensus"
repeat_region     28187..28431      /note="L2 repeat: matches 30..307 of consensus"
repeat_region     29113..29319      /note="L2 repeat: matches 2476..2693 of consensus"
repeat_region     29643..29942      /note="L2 repeat: matches 1..293 of consensus"
repeat_region     30116..30180      /note="L2 repeat: matches 1..406 of consensus"
repeat_region     30181..30485      /note="L2 repeat: matches 1..305 of consensus"
repeat_region     30486..30861      /note="L2 repeat: matches 1..406 of consensus"
repeat_region     30889..31229      /note="L2 repeat: matches 7547..7904 of consensus"
repeat_region     31681..32033      /note="L2 repeat: matches 2379..2709 of consensus"
repeat_region     32163..32593      /note="L2 repeat: matches 1..434 of consensus"
repeat_region     32694..32976      /note="L2 repeat: matches 1..298 of consensus"
repeat_region     33377..33567      /note="L2 repeat: matches 3..201 of consensus"
repeat_region     33583..33734      /note="L2 repeat: matches 9..189 of consensus"
repeat_region     33759..33812      /note="L2 repeat: matches 128..181 of consensus"
repeat_region     33813..34125      /note="L2 repeat: matches 1..296 of consensus"
repeat_region     34126..34234      /note="L2 repeat: matches 21..128 of consensus"
repeat_region     34300..34598      /note="L2 repeat: matches 93..361 of consensus"
repeat_region     34899..35075      /note="L2 repeat: matches 349..525 of consensus"
repeat_region     35076..35177

```

```

repeat_region     35183..35482      /note="L2 repeat: matches 1..108 of consensus"
repeat_region     81989 ACCAGCTTTATATATA 82005
Query Match      94.4%; Score 17; DB 9; Length 175940;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ACCAGCTTTATATATA 17
Db      81989 ACCAGCTTTATATATA 82005

RESULT 8
LOCUS      AY160838
DEFINITION      Uncultured bacterium clone COB P1-13 16S ribosomal RNA gene,
ACCESSION      AY160838
VERSION      AY160838.1 GI:372222277
KEYWORDS
SOURCE      uncultured bacterium
ORGANISM      Bacteria; environmental samples.
REFERENCE      1 (bases 1 to 719)
AUTHORS      Schmitt-Wagner D., Friedrich, M.W., Wagner, B. and Brune, A.
TITLE      Phylogenetic Diversity, Abundance, and Axial Distribution of
            Bacteria in the Intestinal Tract of Two Soil-Feeding Termites
            (Cubitermes spp.)
            Appl. Environ. Microbiol. 69 (10), 6007-6017 (2003)
REFERENCE      2 (bases 1 to 719)
AUTHORS      Schmitt-Wagner D., Friedrich, M.W., Wagner, B. and Brune, A.
TITLE      Direct Submission
JOURNAL      Submitted (09-OCT-2002) Fachbereich Biologie, Mikrobielle
            Oekologie, Universitaet Konstanz, Universitaetstr. 10, Konstanz
            D-78457, Germany
FEATURES
source
1..719
/organism="uncultured bacterium"
/mol_type="genomic DNA"
/isolation_source="termite intestinal tract"
/species="Cubitermes orthognathus, soil-feeding
termite"
/db_xref="taxon:77133"
/clone="COB P1-13"
/environmental_sample
<1..>719
/product="16S ribosomal RNA"

ORIGIN
Query Match      88.9%; Score 16; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 1,3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 CAGCTTTATATATAG 18
Db      49 CAGCTTTATATATAG 34

RESULT 9
LOCUS      AB018110/c
DEFINITION      Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K19M13.
ACCESSION      AB018110 BA000015
VERSION      AB018110.1 GI:3702728
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE      1
AUTHORS      Kaneo, T., Katoh, T., Sato, S., Nakamura, Y., Asamizu, E., Kotani, H.,

```

TITLE Miyajima, N. and Tabata, S.  
Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
Sequence features of the regions of 1,011,550 bp covered by  
sevenfold P1 and TAC clones  
JOURNAL DNA Res. 6 (3), 183-195 (1999)  
MEDLINE 99397451  
PUBMED 10470850  
REFERENCE 2 (bases 1 to 42563)  
AUTHORS Nakamura, Y.  
JOURNAL Direct Submission  
Submitted (06-OCT-1998) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research, 153-2-3, Yata,  
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,  
Tel:81-438-52-3935, Fax:81-438-52-3934)  
Addresses for correspondence: kao@kazusa.or.jp  
For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/agg\_graph.cgi?c=K19M13  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.  
The software programs used to predict genes include: Graal  
(Informatics Group, Oak Ridge National Laboratory,  
http://combio.ornl.gov/Graal-1.3/),  
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of  
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
SolicPredictor (Volker Brendel, Stanford University,  
http://genemini.zozi.iastate.edu/cgi-bin/sp.cgi).  
Genes encoding tRNAs are predicted by tRNAscan-SE  
(Sean Eddy, Washington University School of Medicine, St. Louis,  
http://genome.wustl.edu/eddy/tRNAscan-SE/).  
This sequence may not be the entire insert of this clone. It may be  
shorter because we remove overlaps between neighboring submissions.  
The 5' clone is T33624 and the 3' clone is MQM1.

FEATURES  
Source  
1..42563  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="K19M13"  
/clone\_1lb="Mitsui TAC"  
/ecotype="Columbia"  
2217..3986  
/note="gene\_id:K19M13.1"  
/codon\_start=1  
/evidence=not\_experimental  
/product="disease resistance protein-like"  
/protein\_id="BAB09556.1"  
/db\_xref="GI:9759078"  
/translation="MONUKVYIAMLNLFFSALVRNVLSSGOVICSODRATLLGFKS  
SIEEDTVGIDSWGKDCNGDMGVCNPATNGVGTGVLQSAVNEPILVKKGTLSPS  
LGNIRSLIELITGKRTFTGSIPIPSFNUITSLRQIIDDNSLQCNVSLSLHPLLEI  
LSLGNRPSGLVLPASFGSLRLTLTNLAKRNSFGSIPVTFKNLLELNLDSNLISG  
PIPDIGQFQNLINLIVLSNRSFGLPVSVYSLRQLQMSLERNGLTGPLDSRYSYK  
SLTSLQGNKFIPIASITGLQMLNSRLNLSFDPPLPVGARGPPLSLDISY  
NNLNLGAIPSWIRDKSLDNLACGKRGTPKLTPTTLTSLDLSLDFLGVSAPL  
TSLTNVQKWIIRKSKOLRPDLSKTLPGVASTLDSLRNTLSLNNKSSPLREI  
HUTNQISGRIPDREGSINLKVNLNIGSKIGQIPSSISNLVELYRLDISKNTTGGI  
PQALGQAQLKWLDSLIALTGRIPLDSLINTKTKIKNASFRANRLCGQIPQGRPFNIFP  
AAAYLHNLCICGKPLPACRKTGM"  
complement(join(7408..7601,7792..7852))  
/note="unnamed protein product; gene\_id:K19M13.3  
unknown protein"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB09557.1"  
/db\_xref="GI:9759079"  
/translation="VWLLGGLDWTKEHMSVEDPSKXKSWRLIANVPGKPKLAWG  
HSTCVGSTRVLVIGHNIGEMILNHEHLCLASQSDSL"  
complement(join(10326..10490,10573..10656,10748..10792,  
10888..10942,11028..11147,11224..11313,11604..11725))  
/note="unnamed protein product; gene\_id:K19M13.4

unknown protein"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB09558.1"  
/db\_xref="GI:9759080"  
/translation="MAGPSTNSAPKORRVEATSSNTSTTLRAKDSAPALCLV  
FGCCCCSQAQVETQAEAKKPAEKKTTSQGPRLKKTNDKSKSSINRKR  
PLTAPFLTMSDFRKTFFSKHNSGLAKDAATIGEEKSLTEERKYLIDKALELKEY  
NKSLESNDADDEEEDKOSDVEDDAEKKVDODDEVEERKVENTDKDEKKEE  
EERLIDY"  
complement(join(15556..15709,15838..15906,16013..16129,  
16302..16414,16499..16573,16898..17032,17303..17559,  
17800..18312,18408..18990,19164..19238,19316..19386,  
19825..19927,20466..20692,20785..20921,20999..21140,  
21449..21476))  
/note="unnamed protein product; contains similarity to  
unknown protein  
gb|AA01474.1  
gene\_id:K19M13.6"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB09559.1"  
/db\_xref="GI:9759081"  
/translation="WTKRXYKIOEPFAHSAVNCIKIGKRSRVLYTGEDEKNNLM  
AIGKPNALIVSSRVLCCOSLYGHSSGIDSVTADSEVLAAGAAGTILKNDLEAK  
IVRLTCHRNCSISVDPHPGEPFASGLDNTNLIKIDIRKGCITHYKGTGVNVL  
ETPDGRVNVGGEEDNIYKWDLTAGKLTLEFKHEGOIQLPFPHEPFLATGADRT  
VKFMDLTFELISGGPETAQVRCISPNPGKTVLGLDLSKIFEMEPRIQGDVY  
GSRSLPMNHFLKGLCCSYNOGCVGVVVDLSVPLSPLSFFCLIVGNSLYEPTH  
FQHGQETECMAGDTLOSGNHBEKRSQSRDPVYVLDNNSKTVLAKLSVQNDPL  
KETSILRLSIVQNSDSTYKTSIGSSTVSQSSSKSRPLGLSVQNSDVKE  
SRTSSSTGSLPGTPHVSNTVSKATSGTAVSNATSRNPTKANPKNVAKAD  
FAPYIVRAPRPIEQAATESRAELDIARTPYLQADSRSSSNNDPLPASYLE  
MSEQAPVEPNPIPFGTLPFGKGMKATERSINDPRYKXGRSRSRMSGSPND  
ENVYLVSRGNRDPSPTEQGGDSQSLQSLFKMKRISFMCLISGPFQSLVIRE  
RGRFRSPBPVNSFSGNMPAPRIYRSMNFKRGKHMPEQSIDSPSENYEDIMG  
KKNQVSSWQSRKLKQVRRYWRNDVKSISIKHMADNATVADVLITERNL  
TLNDCTSLPLTALGSGMDQHLVSLDLTLVLYSGPIYSLSAPASVGDIEA  
EORIERYSRCFVELKEVKACLPISLISGSIITWTSLTLDKNAQKQVFLVEEPAS  
SOVNHLLPFVDLVPKGV"  
complement(24844..25392)  
/note="unnamed protein product; contains similarity to  
ferredoxin-thioredoxin reductase variable chain  
gene\_id:K19M13.7"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB09560.1"  
/db\_xref="GI:9759082"  
/translation="WSSQIALSPAIAAIRPSSHDCLSASATTAATPMALKCTIVA  
PLSLFTSQSQIKHSSSRKTRTTRICDVALTKSDSINADNPSSPSEEEIEAEKA  
KIGSRVVTAPLKYHVNRYPEVDLEMEGKLADVYAVWGXKISINLPIKEFFKEI  
EGRLVYFVHLKEDDEPFDQ"  
complement(join(26890..27881,28441..28540,28657..28803,  
29264..29301,29417..29510,29664..30038,30122..30574))  
/note="unnamed protein product; gene\_id:K19M13.8  
unknown protein"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB09561.1"  
/db\_xref="GI:9759083"  
/translation="WQKSGVNRNPSLKVAILPQAQSLRLRIGCSQIATGSGQSSPIV  
FPEKRNKVVASRRGEVNDPQKPDGRHIDIGGDEKSLDLSLVLVAGLVLDK  
RKSASGDATIEOOPATDLSIKKAVDAKLTSLALVSGMDQLNDVSVTVVGLRH  
FTVAYVPIGSGSGLSGFTKPKRSRGRFVAPTVBEAVQVWSPDDOQFCINCLP  
LVAKKQASSLPSVPIPTPEPLVFRCKSAAPKVIINPSSGGRSIVKPNVVEPIEX  
LAGIKMEVNTKAGARELASTVDINLSDGIIICGGDIIIEVNLGLITNSNPKEG  
VSLPIGIVPAGSDNSLWTVLGVADPISALSLVKGSLATVDFAVEVHLHTGIIHGM  
TVSYGFGVSVLELSEKYGQRFGLRFLVGFGLFNCPLPYTSYVLEVLPOKDEAGX  
IRLEKEVNDQDLYTDWRRSREGFPRLASISISIMPSVSGEIDTSSTIATSEP  
SEVYRGIDPMKRLSGSRDVTAPREHIVQAOQSTPNMRPSSKSRMKQGMGLSV  
ODPPTRSKMGNTGGQDREDSISVSDPGPIWMDGPKMDPERSAWVDENSIETLDPGED  
IETLRQOSTIPIEDFWVRKGFGLINVCNACKCTVQSSQVAVANSHDCTMDML  
LVHGCGRLRLRPFILLQTRHSLSPVECCVKNVLSYFLTLNLIQLAM"



CDS  
34686..35021  
/note="unnamed protein product; gene\_id:K19M13.9  
/unknown protein  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB09562.1"  
/db\_xref="GI:9759084"  
/translation="MVFEGGFYINLVGASDASPKLALIGETLKQFPYSNPEETRC  
GKLRDRHREGRFTSPQSRLLTVADRNFCGVPVNPRLNMTKRLTSLHLVDTKS  
TCVSCSLSF"  
35475..36383  
/note="unnamed protein product; gb|AAC78512.1  
gene\_id:K19M13.10  
strong similarity to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB09563.1"  
/db\_xref="GI:9759085"  
/translation="MSLVENLFLFEPQSKRKINPFTIESPNCNMFVNRKICISCY  
TVDNFEGSFYLYKGMQNSKALFTKGLISQTSMLDEKHLVLDLDTIHLITIK  
TSLVSEKSLIEVESRKDIKRFNTGFPESLIKLRFVHQLKCNEMFSMTYIK  
GGDYARLVLEMIDPDKFFGNRVITRRESFGKTLDLVADRGIVIVDTTSVMPH  
DKNLIOIARYKXFKGSKCLFSEDKKIDSDERKPLMTALRFKDVHEEFYDMSKK  
DLDSVVRPLKELSLRWKKAQGV"  
join(37928..38528,38657..38766,39049..39108,39289..39339,  
39675..40085,40152..40239,40345..40769)  
/note="unnamed protein product; emb|CAB62634.1  
gene\_id:K19M13.11  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB09564.1"  
/db\_xref="GI:9759086"  
/translation="MEDNKEKVGSSKKRLRKPKSLFVGMGSRNLIIEPLSGRDT  
NKISENDVTAIIIMVYREKRETPKSKRKRTVACDEKRLLEFGAKIVIKPDLV  
EKIYVQOEEDLPRTDIIYASDDKQQLSLSDYAKOTKQVSKRGTPFAIYRTYK  
LVLRKSLVQELATPFTESKVRITRKINPQVAVTVGKSGDIDGNLQVITMY  
SVYIKDVTSLSDDDFQSECEELHQRINNGFAKRLTVDMEKASLSHEDVKHL  
EKRELQNPDEQKLVDEPEIYVAEELLEPECEDDDDRTIEDSLIVNPAHOSDKQ  
RORDLPVSSSVKSKQENSLKPNBEOLRLCDVPDVAEKLPEFVDDQKLVNDVP  
YQTKPSLRIRIKWKYKNQVRPLICNNSIGLKLGLANPRLCQSNKDLHEDEV  
EPNANGITLMTDSTTBSEMNTRKSKHOSSTPVIDLSKTKQAHSPRIIIEISDDDD

Query Match 88.9%; Score 16; DB 8; Length 42563;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAGCTTTTATTA 17  
|||||  
Db 33745 CCAGCTTTTATTA 33730

RESULT 10  
AF467900/c 48843 bp DNA linear PLN 22-MAY-2003  
LOCUS Prunus persica clone pbn31C7 hypothetical protein, and  
DEFINITION hypothetical transcription factor genes, complete cds.  
ACCESSION AF467900  
VERSION AF467900.1 GI:27450528  
KEYWORDS  
SOURCE  
ORGANISM  
Prunus persica (peach)  
Prunus persica  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.  
1 (baaes 1 to 48843)  
Georgi, L., Wang, Y., Yvergnaux, D., Ormsbee, T., Inigo, M.,  
Reignard, G. and Abbott, G.  
Construction of a BAC library and its application to the  
identification of simple sequence repeats in peach [Prunus persica  
(L.) Batsch]  
Theor. Appl. Genet. 105 (8), 1151-1158 (2002)  
JOURNAL  
PUBMED  
12582893

REFERENCE 2 (baaes 1 to 48843)  
AUTHORS Georgi, L., Wang, Y. and Abbott, A. G.  
TITLE Direct Submision  
JOURNAL Submitted (11-JAN-2002) Genetics and Biochemistry, Clemson  
University, 122 Long Hall, Clemson, SC 29634, USA  
FEATURES  
source  
1..48843  
/organism="Prunus persica"  
/mol\_type="genomic DNA"  
/culivar="hemared"  
/db\_xref="taxon:3760"  
/clone="pbn31C7"  
/issue\_type="rootstock"  
/clone\_id="pbelobAC11"  
/note="derived from genomic DNA partially digested with  
Hind III"  
291..318  
/note="similar to a direct repeat on Arabidopsis thaliana  
chromosome 2"  
/tpc\_type="inverted"  
355..393  
/note="similar to a direct repeat on Arabidopsis thaliana  
chromosome 2"  
/tpc\_type="inverted"  
join(3833..3907,4079..4288,4418..4542,4630..5560)  
/product="hypothetical protein"  
join(3833..3907,4079..4288,4418..4542,4630..5560)  
/note="similar to hypothetical protein F6P9.9 in  
Arabidopsis thaliana"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA014624.1"  
/db\_xref="GI:27450529"  
/translation="MGMGRNALICHTVYSRSLIKISVQLFLRESPOVGLASAD  
HLQAKSWLSHSGTSDILPEFGAAYSANQIINVSOIPLIKRCPPRYVLTFTW  
QVCGEESKEVEIIONREIRLEAVYRPSALPLNSVLADVDSGHNDVQPVLPIT  
SIDDEGVDTSSASMAPNIIHTSOSPLSVQIIAASQSVVNNRNHPSNRVAGI  
PUGVEDVYAAASAAAGAIYNSNHNMLIDELIKLSNPKMIEKLVKQPOMTREP  
QNSIADPPPHNIRNTSSSTRSSSTALSGHFYQPVAVAGRPDPAPPAHAISSVS  
PPVAGPPADIVYKSLIQGGDGHDSFPFGNRRSHSALINQSNNGYKSDSRP  
KIMKPIYFNSSRGCGANCAVOHASFQPRSSSAPEVHSTRKMDREISS"  
join(46350..6706,7104..7133)  
/product="hypothetical protein"  
join(6350..6706,7104..7133)  
/note="similar to unknown protein MJ24.9 in Arabidopsis  
thaliana"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA014625.1"  
/db\_xref="GI:27450530"  
/translation="MLKTLRVRIEPRNADPRTASCLEFLAOCNARAKXENSPSCAVO  
VKRRTDEPKRTVTVFNGVEEVDATATPAODIRNIIIEKQSLTEIQMFRDAGEFM  
PVIIIPAEELHQPAGTKPRKADSKQ"  
complement(join(47784..7882,7973..8098,8313..8512,  
8627..8865,9253..9440,9712..10323,10419..10574))  
/product="hypothetical protein"  
complement(join(7784..7882,7973..8098,8313..8512,  
8627..8865,9253..9440,9712..10323,10419..10574))  
/note="similar to unknown protein MJ24.8 in Arabidopsis  
thaliana"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA014626.1"  
/db\_xref="GI:27450531"

misc\_feature  
join(12687..15371,15371..15549,16324..16402,16405..16717)  
/note="similar to putative DEAH helicase MPE16.9 in Arabidopsis thaliana; likely non-functional due to two frame-shifts and premature stop codons"  
<17499..>18764  
/product="hypothetical protein"  
17499..18764  
/note="similar to unknown protein K17E12.5 in Arabidopsis thaliana"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA014627.1"  
/db\_xref="GI:27450532"  
/translation="MGSVSLKIGDGTAFKRAATLCSAVMLMLFSVITTNLPALVAF  
TSPPQOQTHLHHTQKNSLISEVSLIREDSOKLAQKELGYESIDSR  
SNVAHEKLFLOHLOPLGKDSRGITGEMASVDSKSDLLSOVMYKSGPCD  
DMSLAQKILRGCEPIPRRCRAKTLPKVGLNPPRILMKPVSGKLTITGSGKSE  
CLNSKLSRDCVCGFDLVNGENORFYKAKNDPLIDVLAISGGGIIIGFDIGGS  
GTPAARVAERNMTVITNTLIDAPSEFIFARGLPFLSLDPPYDVNFDLVHAA  
SGLDVGGKPKPEFEMDILRLPGGLFMDNFYCSNEKKRDLTLIERFGYKLLK  
WVVDKVDAAAGSKSEVYLAVLQKPVRA"  
complement(join(19604..19745,19833..20023,20115..21436,  
21513..21648,21782..21857,22120..22242,22363..22527,  
22823..22913,23088..23172,23278..23430,23534..23590,  
23701..23799,24052..24164,24372..24483))  
/product="hypothetical transcription factor"  
complement(join(19604..19745,19833..20023,20115..21436,  
21513..21648,21782..21857,22120..22242,22363..22527,  
22823..22913,23088..23172,23278..23430,23534..23590,  
23701..23799,24052..24164,24372..24483))  
/note="similar to IAA24/monopteros/ARF5 gene, FeF9.10  
protein, in Arabidopsis thaliana"  
/codon\_start=1  
/product="hypothetical transcription factor"  
/protein\_id="AA014628.1"  
/db\_xref="GI:27450533"  
/translation="MINGSVEEIKKAGGLSGAQSILDEKELKELOHSGSKRAIN  
SELMACAGLVLCIPVGSLSYFPQHSQAVSTKRTATISQIPNPNLPSOLQV  
ONVTLHADKETDEIYAQMSLKPVNSEKDPVFPFGKPKSGHPEFCKTLTADST  
HGGSFVRAAEKLFPLDPTMOPRSGEIVRLDHSNMFRIYRROPKRLITTM  
SLFVGAKRLLAGSGLFIRDEKSLMGRRANQOTTLPSVSASMIHGVAAAA  
HAANSPFTIFNPRACPSFEVPLATYQALITGLOSVMRGMPFEESKRY  
MGITVSTSDLPRLPSPSKWNLQVDEPCCQKORNSWELETENETIFPFLTS  
SLRPSHTGFLGATEWGNLIRKPIFVPEIGNNSFPYSISNLSQOLVMLKPOL  
VNAAGTALAQOOSPPANGDLADKMAQKLIQKNPVSGTSLCSQONPOSSLDSS  
ATIDVNTTSHAILPKLNLTKRGSQAPVGNSTDTKLETFDSADLSQNSGTGIE  
DKAAGVSPYNLQUTRANQSSAQLQTSPPMPPLSILYHQQDPMPSDN  
STNGSLFNDNCTFYQSTQFPAKTRSGQSVFGLQSSAVLITENNSSLTISIG  
EMWNSLNLNRLLPQVDQLTSHQGQSLNCSNSSLRLSDSSNQSGIYGGPND  
VSGSVAVIDPVSSTILDFSTLKNADFNHPSDDLGNLSSQDLQSLTASIGDS  
QAFSRDLADNSGSSSNIDLDESSLONNNSHWQVPPRYTQKQTSVRSID  
VPSFKVVEELCSAIECMRGLGLNDPRGSGMKLYVVDYENDVLVDDPMEERVGVY  
RCRILSPTEVQMSSEGIKLNANMGINGTMSBGRV"  
join(36926..36970,36973..37377)  
/note="similar to hypothetical protein FeF9.11 in Arabidopsis thaliana; non-functional due to a frame-shift"  
41043..41144  
repeat\_region  
41144..41437  
/note="likely the remains of a retroelement"  
repeat\_region  
41148..42253  
/note="likely the remains of a retroelement"  
repeat\_region  
42255..42410  
/note="likely the remains of a retroelement"

Db 5275 CAAGCTTTATATAG 5260  
RESULT 11  
AC145628  
LOCUS  
DEFINITION  
Homo sapiens chromosome UNK clone XAFOS-802760B8, \*\*\* SEQUENCING IN PROGRESS \*\*\*; 8 unordered pieces.  
ACCESSION  
AC145628  
VERSION  
AC145628.1 GI:32996967  
KEYWORDS  
HTG: HTGS PHASE1.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 51797)  
Wilson, R.K.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 51797)  
Wilson, R.K.  
Direct Submission  
Submitted (19-JUL-2003) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA  
COMMENT  
----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_AA802760B8  
----- Summary Statistics -----  
Sequencing vector: plasmid; 100%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator B3g Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 42380 bases at least Q40  
Consensus quality: 44639 bases at least Q30  
Consensus quality: 46559 bases at least Q20  
-----  
NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
1 3260: contig of 3260 bp in length  
3261 3360: gap of unknown length  
3361 7532: contig of 4172 bp in length  
7533 7632: gap of unknown length  
7633 12296: contig of 4666 bp in length  
12296 12398: gap of unknown length  
12398 21234: contig of 8836 bp in length  
21234 21334: gap of unknown length  
21334 46167: contig of 24833 bp in length  
46167 46267: gap of unknown length  
46267 47548: contig of 1281 bp in length  
47548 47648: gap of unknown length  
47648 49886: contig of 2238 bp in length  
49886 49987: gap of unknown length  
49987 51797: contig of 1811 bp in length.  
FEATURES  
Source  
1. 51797  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="UNK"  
/clone="XAFOS-802760B8"  
misc\_feature 1..3260

Query Match 88.9%; Score 16; DB 8; Length 48843;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3 CAAGCTTTATATAG 18  
|||||

misc\_feature

1..3260



```

misc_feature /note="assembly_name:Contig10"
3361..7532
/note="assembly_name:Contig11"
7633..12298
misc_feature /note="assembly_name:Contig12"
12399..21234
misc_feature /note="assembly_name:Contig13"
21335..46167
/note="assembly_name:Contig14"
46268..47548
misc_feature /note="assembly_name:Contig6"
47649..49886
/note="assembly_name:Contig8"
49987..51797
misc_feature /note="assembly_name:Contig9"

```

## ORIGIN

```

Query Match      88.9%; Score 16; DB 2; Length 51797;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3  CAACTTTTAAATAG 18
Db      42039 CAACTTTTAAATAG 42054

```

## RESULT 12

```

AC109350      58068 bp  DNA  linear  PRI 30-MAR-2002
LOCUS
DEFINITION Homo sapiens BAC clone RP11-267D2 from 4, complete sequence.
ACCESSION AC109350
VERSION AC109350.5 GI:19526559
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PMID 9847074

```

```

REFERENCE
AUTHORS Buatesi, D. and Cotton, M.
TITLE The sequence of Homo sapiens BAC clone RP11-267D2
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

REFERENCE
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 58068)
REFERENCE
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 58068)

```

```

REFERENCE
AUTHORS Direct Submission
JOURNAL Submitted (30-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 17, 2002 this sequence version replaced gi:19339131.
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center

```

```

Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0267D02
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, F.Y., Zhao, B., Frengen, E., Tateno, M., Cataneese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-609N18, 2000 bp overlap; the clone sequenced to the right is CTD-2012117, 2000 bp overlap. Actual end of this clone is at base position 105788 of CTD-2012117.

Polymorphisms exist between AC093722 and AC109350.

## FEATURES

## SOURCE

```

1..58068
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-267D2"
/clone_lib="RPCT-11"
423..462
/rpc_family="AT_rich"
1124..1158
/rpc_family="AT_rich"
1283..1310
/rpc_family="AT_rich"
1753..1809
/rpc_family="AT_rich"
2483..2513
/rpc_family="AT_rich"
3343..3373
/rpc_family="AT_rich"
3694..3714
/rpc_family="AT_rich"
5020..5146
/rpc_family="AT_rich"
5209..5251
/rpc_family="MIR"
5524..5626
/rpc_family="(CATR)n"
5666..5954
/rpc_family="Alu"

```



|           |  |
|-----------|--|
| ACCESSION | SEQUENCE, 3 ordered pieces.  |
| VERSION   | AC092359   |
| KEYWORDS  | AC092359.2 GI:19224822   |
| SOURCE    | HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.  |
| ORGANISM  | Homo sapiens (human)   |
| REFERENCE | Homo sapiens   |
| AUTHORS   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.          |
| TITLE     | 1 (bases 1 to 107129)  |
| JOURNAL   | DOE Joint Genome Institute.  |
| REFERENCE | Sequencing of Human Chromosome 5   |
| AUTHORS   | Unpublished  |
| TITLE     | 2 (bases 1 to 107129)  |
| JOURNAL   | DOE Joint Genome Institute.  |
| REFERENCE | Direct Submission  |
| AUTHORS   | Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA |
| TITLE     | 3 (bases 1 to 107129)  |
| JOURNAL   | DOE Joint Genome Institute.  |
| REFERENCE | Direct Submission  |
| AUTHORS   | Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA |
| TITLE     | On Mar 7, 2002 this sequence version replaced gi:14589548.   |
| JOURNAL   |  |
| COMMENT   |  |

```
Project Information
Center Project Name: 537429
Center clone name: RPCI-11_341E24
```

```

Summary Statistics
Consensus quality: 104726 bases at least Q40
Consensus quality: 105971 bases at least Q30
Consensus quality: 106638 bases at least Q20
Estimated insert size: 181520; agarose-fp estimation
Estimated insert size: 166939; sum-of-contigs estimation
Quality coverage: 6.12 in Q20 bases; agarose-fp estimation
Quality coverage: 10.39 in Q20 bases; sum-of-contigs estimation
NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the subitroor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 27924: contig of 27924 bp in length
* 27925 28024: gap of unknown length
* 28025 29265: contig of 1245 bp in length
* 29270 29365: gap of unknown length
* 29370 107129: contig of 77760 bp in length.

```

```
FEATURES
  source
    1. .107129
      Location/Qualifiers
        location="Homo sapiens"
        mol_type="genomic DNA"
        db_xref="taxon:9606"
        chromosome="5"
        clone="RP11-341E24"
        clone_11b="RP11 human BAC library 11"
```

|                       |                 |               |           |                |
|-----------------------|-----------------|---------------|-----------|----------------|
| Query Match           | 88.9%           | Score 16;     | DB 2;     | Length 107129; |
| Best Local Similarity | 100.0%          | Pred. No. 74; |           |                |
| Matches 16;           | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0;        |

QY 3 CAAGCTTTTATAATAG 18  
|||||  
Db 68444 CAAGCTTTTATAATAG 68429

|            |  |
|------------|--|
| RESULT 15  |  |
| LOCUS      | ACI04793/c   |
| DEFINITION | Homo sapiens BAC clone RP11-23406 from 4, complete sequence.   |
| ACCESSION  | ACI04793   |
| VERSION    | ACI04793.4   |
| KEYWORDS   | GI:19909440  |
| SOURCE     | HTG.   |
| ORGANISM   | Homo sapiens (human)   |
|            | Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.   |
| REFERENCE  |  |
| AUTHORS    | Gulston,J.E. and Waterston,R.  |
| TITLE      | 1 (bases 1 to 133244)  |
| JOURNAL    | Toward a complete human genome sequence  |
| MEDLINE    | Genome Res. 8 (11), 1097-1108 (1998)   |
| PUBMED     | 99063792   |
| REFERENCE  |  |
| AUTHORS    | 2 (bases 1 to 133244)  |
| TITLE      | Swarczenegen-Shahid,S., Kozlowski,A. and Dignan,G.   |
| JOURNAL    | The sequence of Homo sapiens BAC clone RP11-23406<br>Unpublished (2001)  |
| REFERENCE  |  |
| AUTHORS    | 3 (bases 1 to 133244)  |
| TITLE      | Waterston,R.H.   |
| JOURNAL    | Direct Submission<br>Submitted (21-DEC-2001) Genome Sequencing Center, Washington<br>University School of Medicine, 4444 Forest Park Parkway, St. Louis,<br>MO 63108, USA  |
| REFERENCE  |  |
| AUTHORS    | 4 (bases 1 to 133244)  |
| TITLE      | Waterston,R.H.   |
| JOURNAL    | Direct Submission<br>Submitted (11-MAR-2002) Genome Sequencing Center, Washington<br>University School of Medicine, 4444 Forest Park Parkway, St. Louis,<br>MO 63108, USA  |
| REFERENCE  |  |
| AUTHORS    | 5 (bases 1 to 133244)  |
| TITLE      | Waterston,R.H.   |
| JOURNAL    | Direct Submission<br>Submitted (03-APR-2002) Genome Sequencing Center, Washington<br>University School of Medicine, 4444 Forest Park Parkway, St. Louis,<br>MO 63108, USA  |
| REFERENCE  |  |
| AUTHORS    | 6 (bases 1 to 133244)  |
| TITLE      | Waterston,R.   |
| JOURNAL    | Direct Submission<br>Submitted (16-APR-2002) Department of Genetics, Washington<br>University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA<br>On Apr 3, 2002 this sequence version replaced gi:19339126. |
| COMMENT    |  |

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:** Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Catanesi, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the right is RP11-80B23, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-23406; actual end is at base position 3275 of RP11-80B23.

The region from 17296 to 17363 is covered only by a PCR product from clone DNA.

```

FEATURES
    source
        Location/Qualifiers
            1..133244
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="4"
                /map="4"
                /clone="RP11-23406"
                /clone_id="RPCI-11"
                1..275
                    repeat_region
                        /rpt_family="MaLR"
                        899..977
                            repeat_region
                                /rpt_family="L2"
                                2465..2496
                                    repeat_region
                                        /rpt_family="AT_rich"
                                        2506..2709
                                            repeat_region
                                                /rpt_family="ERV_L"
                                                2710..3136
                                                    repeat_region
                                                        /rpt_family="MaLR"
                                                        3137..3372
                                                            repeat_region
                                                                /rpt_family="ERV_L"
                                                                3794..3923
                                                                    repeat_region
                                                                        /rpt_family="MIR"
                                                                        4072..4222
                                                                            repeat_region
                                                                                /rpt_family="ERV_L"
                                                                                4274..4710
                                                                                    repeat_region
                                                                                        /rpt_family="ERV_L"
                                                                                        5294..5377
                                                                                            repeat_region
                                                                                                /rpt_family="ERV_L"
                                                                                                5477..5732
                                                                                                    repeat_region
                                                                                                        /rpt_family="ERV_L"
                                                                                                        5953..6636
                                                                                                            repeat_region
                                                                                                                /rpt_family="ERV_L"
                                                                                                                6726..7020
                                                                                                                    repeat_region
                                                                                                                        /rpt_family="Alu"
                                                                                                                        8017..8117
                                                                                                                            repeat_region
                                                                                                                                /rpt_family="(TA)n"
                                                                                                                                8599..8631
                                                                                                                                    repeat_region
                                                                                                                                        /rpt_family="AT_rich"
                                                                                                                                        8674..8924
                                                                                                                                            repeat_region
                                                                                                                                                /rpt_family="ERV_L"
                                                                                                                                                8931..9806
                                                                                                                                                    repeat_region
                                                                                                                                                        /rpt_family="ERV_L"
                                                                                                                                                        9791..9891
                                                                                                                                                            repeat_region
                                                                                                                    /rpt_family="ERV_L"
                                                                                                                    9951..10767
                                                                                                                        repeat_region
                                                                                                                            /rpt_family="ERV_L"
                                                                                                                            10769..11282
                                                                                                                                repeat_region
                                                                                                                                    /rpt_family="L1"
                                                                                                                                    11288..11594
                                                                                                                                        repeat_region
                                                                                                                                            /rpt_family="Alu"
                                                                                                                                            11597..11729
                                                                                                                                                repeat_region
                                                                                                                                                    /rpt_family="ERV1"
                                                                                                                                                    11866..11893
                                                                                                                                                        repeat_region
                                                                                                                                                            /rpt_family="(TA)n"
                                                                                                                                                            11901..12013

```

```

repeat_region
    /rpt_family="ERV1"
    12853..12922
repeat_region
    /rpt_family="ERV1"
    12923..13229
repeat_region
    /rpt_family="Alu"
    13330..13606
repeat_region
    /rpt_family="ERV1"
    14053..14471
repeat_region
    /rpt_family="MaLR"
    14537..15128
repeat_region
    /rpt_family="ERV1"
    15255..15553
repeat_region
    /rpt_family="Alu"
    15723..15743
repeat_region
    /rpt_family="AT_rich"
    15987..16007
repeat_region
    /rpt_family="AT_rich"
    16386..16548
repeat_region
    /rpt_family="MIR"
    16557..17146
repeat_region
    /rpt_family="ERV_L"
    17170..17363
repeat_region
    /rpt_family="(TTATG)n"
    17365..17498
repeat_region
    /rpt_family="Alu"
    18610..18796
repeat_region
    /rpt_family="MIR"
    18801..18823
repeat_region
    /rpt_family="AT_rich"
    20966..21013
repeat_region
    /rpt_family="AT_rich"
    21403..21438
repeat_region
    /rpt_family="(TTTTTG)n"
    22435..24301
repeat_region
    /rpt_family="L1"
    24295..24540
repeat_region
    /rpt_family="L1"
    24621..24910
repeat_region
    /rpt_family="Alu"
    25338..25371
repeat_region
    /rpt_family="(TG)n"
    26548..26849
repeat_region
    /rpt_family="Alu"
    27041..27080
repeat_region
    /rpt_family="AT_rich"
    27135..28394
repeat_region
    /rpt_family="L2"
    28401..28761
repeat_region
    /rpt_family="MaLR"
    28781..29106
repeat_region
    /rpt_family="L2"
    30336..30708
repeat_region
    /rpt_family="MaLR"
    31100..31201
repeat_region
    /rpt_family="MIR"
    31600..31770

```

```

Query March 88.9%; Score 16; DB 9; Length 133244;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CAAGCTTTTATATATG 18
DB 111702 CAAGCTTTTATATATG 111687

```

Search completed: February 10, 2005, 13:14:00  
Job time : 117.027 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 5.5316 Seconds  
(Without alignments)  
5324.730 Million cell updates/sec

Title: US-10-790-430-12

Sequence: 1 accaagctttcataatag 18

Scoring table: OLIGO NUC  
Gapop\_60.0, Gapext 60.0

Searched: 1202784 seqe, 818138359 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Patente NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                | Description        |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1          | 18    | 100.0       | 18     | US-09-872-051-12     | Sequence 12, Appl  |
| 2          | 18    | 100.0       | 1183   | US-09-872-051-8      | Sequence 8, Appl   |
| 3          | 17    | 94.4        | 45587  | US-09-949-016-15836  | Sequence 15836, A  |
| 4          | 15    | 83.3        | 601    | US-09-949-016-129275 | Sequence 129275, A |
| 5          | 15    | 83.3        | 601    | US-09-949-016-129346 | Sequence 129346, A |
| 6          | 15    | 83.3        | 601    | US-09-949-016-129417 | Sequence 129417, A |
| 7          | 15    | 83.3        | 601    | US-09-949-016-129488 | Sequence 129488, A |
| 8          | 15    | 83.3        | 601    | US-09-949-016-129559 | Sequence 129559, A |
| 9          | 15    | 83.3        | 601    | US-09-949-016-129630 | Sequence 129630, A |
| 10         | 15    | 83.3        | 601    | US-09-949-016-129701 | Sequence 129701, A |
| 11         | 15    | 83.3        | 601    | US-09-949-016-129772 | Sequence 129772, A |
| 12         | 15    | 83.3        | 601    | US-09-949-016-129843 | Sequence 129843, A |
| 13         | 15    | 83.3        | 601    | US-09-949-016-129914 | Sequence 129914, A |
| 14         | 15    | 83.3        | 601    | US-09-949-016-129985 | Sequence 129985, A |
| 15         | 15    | 83.3        | 601    | US-09-949-016-130056 | Sequence 130056, A |
| 16         | 15    | 83.3        | 2152   | US-09-449-632-3      | Sequence 3, Appl   |
| 17         | 15    | 83.3        | 107329 | US-09-949-016-12663  | Sequence 12663, A  |
| 18         | 15    | 83.3        | 107329 | US-09-949-016-12664  | Sequence 12664, A  |
| 19         | 15    | 83.3        | 107330 | US-09-949-016-15408  | Sequence 15408, A  |
| 20         | 15    | 83.3        | 107330 | US-09-949-016-15409  | Sequence 15409, A  |
| 21         | 15    | 83.3        | 107330 | US-09-949-016-15410  | Sequence 15410, A  |
| 22         | 15    | 83.3        | 107330 | US-09-949-016-15411  | Sequence 15411, A  |
| 23         | 15    | 83.3        | 107330 | US-09-949-016-15412  | Sequence 15412, A  |
| 24         | 15    | 83.3        | 107751 | US-09-949-016-15413  | Sequence 15413, A  |
| 25         | 15    | 83.3        | 107751 | US-09-949-016-15414  | Sequence 15414, A  |
| 26         | 15    | 83.3        | 107751 | US-09-949-016-15415  | Sequence 15415, A  |
| 27         | 15    | 83.3        | 107751 | US-09-949-016-15416  | Sequence 15416, A  |

|    |    |      |        |                     |                   |
|----|----|------|--------|---------------------|-------------------|
| 28 | 15 | 83.3 | 107925 | US-09-949-016-11875 | Sequence 11875, A |
| 29 | 15 | 83.3 | 107926 | US-09-949-016-15404 | Sequence 15404, A |
| 30 | 15 | 83.3 | 107926 | US-09-949-016-15405 | Sequence 15405, A |
| 31 | 15 | 83.3 | 107926 | US-09-949-016-15406 | Sequence 15406, A |
| 32 | 15 | 83.3 | 107926 | US-09-949-016-15407 | Sequence 15407, A |
| 33 | 15 | 83.3 | 107926 | US-09-949-016-15408 | Sequence 15408, A |
| 34 | 14 | 77.8 | 507    | US-09-621-976-9838  | Sequence 9838, Ap |
| 35 | 14 | 77.8 | 687    | US-08-956-171E-608  | Sequence 608, App |
| 36 | 14 | 77.8 | 687    | US-08-781-986A-608  | Sequence 608, App |
| 37 | 14 | 77.8 | 822    | US-09-134-001C-1178 | Sequence 268, App |
| 38 | 14 | 77.8 | 999    | US-09-417-251A-13   | Sequence 1178, Ap |
| 39 | 14 | 77.8 | 1126   | US-09-417-251A-13   | Sequence 13, Appl |
| 40 | 14 | 77.8 | 1328   | US-08-290-448A-58   | Sequence 58, Appl |
| 41 | 14 | 77.8 | 1328   | US-08-290-448A-58   | Sequence 58, Appl |
| 42 | 14 | 77.8 | 1328   | US-08-175-069A-58   | Sequence 58, Appl |
| 43 | 14 | 77.8 | 1328   | US-08-461-939B-58   | Sequence 58, Appl |
| 44 | 14 | 77.8 | 1332   | US-08-464-000-58    | Sequence 1420, Ap |
| 45 | 14 | 77.8 | 1349   | US-09-489-039A-1420 | Sequence 73, Appl |

#### ALIGNMENTS

RESULT 1  
US-09-872-051-12  
Sequence 12, Application US/09872051  
Patent No. 6825400  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
TITLE OF INVENTION: Corn Event PV-ZMGT32(hk603) and Composition and Methods for Det  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 18  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(18)  
OTHER INFORMATION: Agrobacterium tumefaciens nos 3' terminator and rice actin prom  
OTHER INFORMATION: et DN  
US-09-872-051-12  
Query Match 100.0%; Score 18; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1  
1 ACCAAGCTTTATATAG 18  
1 ACCAAGCTTTATATAG 18  
RESULT 2  
US-09-872-051-8  
Sequence 8, Application US/09872051  
Patent No. 6825400  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl

```
/ APPLICANT: Hironaka, Catherine
/ APPLICANT: Heck, Gregory
/ APPLICANT: You, Jinsong
/ TITLE OF INVENTION: Corn Event PV-ZMG132 (nk603) and Composition and Methods for Detect
/ TITLE OF INVENTION: Thereof
/ FILE REFERENCE: 38-21(52258)B
/ CURRENT APPLICATION NUMBER: US/09/872.051
/ PRIOR FILING DATE: 2001-06-01
/ PRIOR APPLICATION NUMBER: 60/213,567
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 60/241,215
/ PRIOR FILING DATE: 2000-10-13
/ PRIOR APPLICATION NUMBER: 60/240,014
/ PRIOR FILING DATE: 2000-10-13
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 8
/ LENGTH: 1183
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ NAME/KEY: source
/ LOCATION: (1)..(1183)
/ OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
/ OTHER INFORMATION: 165-381 construct vector DNA
/ OTHER INFORMATION: 382-686 Zea mays plastid genes, rps11 and rpoA
/ OTHER INFORMATION: 687-1183 Zea mays genomic DNA
US-09-872-051-8
```

```
Query Match          100.0%; Score 18; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1  ACCAAGCTTTATATAG 18
Db      156  ACCAAGCTTTATATAG 173
```

```
RESULT 3
US-09-949-016-15836/c
/ Sequence 15836, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949.016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15836
/ LENGTH: 45587
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-15836
```

```
Query Match          94.4%; Score 17; DB 4; Length 45587;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1  ACCAAGCTTTATATA 17
Db      6030  ACCAAGCTTTATATA 6014
```

RESULT 4

```
US-09-949-016-129275
/ Sequence 129275, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949.016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 129275
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-129275
```

```
Query Match          83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2  CCAAGCTTTATAT 16
Db      548  CCAAGCTTTATAT 562
```

```
RESULT 5
US-09-949-016-129346
/ Sequence 129346, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949.016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 129346
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-129346
```

```
Query Match          83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2  CCAAGCTTTATAT 16
Db      548  CCAAGCTTTATAT 562
```

```
RESULT 6
US-09-949-016-129417
/ Sequence 129417, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129417
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129417
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTAAAT 16
      |||||
Db      548 CCAAGCTTTTAAAT 562
```

```
RESULT 7
US-09-949-016-129488
; Sequence 129488, Application US/09949016
```

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129488
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129488
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTAAAT 16
      |||||
Db      548 CCAAGCTTTTAAAT 562
```

```
RESULT 8
US-09-949-016-129559
; Sequence 129559, Application US/09949016
```

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
```

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129559
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129559
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTAAAT 16
      |||||
Db      548 CCAAGCTTTTAAAT 562
```

```
RESULT 9
US-09-949-016-129630
; Sequence 129630, Application US/09949016
```

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129630
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129630
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTAAAT 16
      |||||
Db      548 CCAAGCTTTTAAAT 562
```

```
RESULT 10
US-09-949-016-129701
; Sequence 129701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129701
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129701
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
         |||||
Db       548 CCAAGCTTTTATAT 562
```

```
RESULT 11
US-09-949-016-129772
; Sequence 129772; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129772
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129772
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
         |||||
Db       548 CCAAGCTTTTATAT 562
```

```
RESULT 12
US-09-949-016-129843
; Sequence 129843; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129843
; LENGTH: 601
```

```

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129843
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
         |||||
Db       548 CCAAGCTTTTATAT 562
```

```
RESULT 13
US-09-949-016-129914
; Sequence 129914; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129914
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129914
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
         |||||
Db       548 CCAAGCTTTTATAT 562
```

```
RESULT 14
US-09-949-016-129985
; Sequence 129985; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129985
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-129985
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
```



Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 CCAAGCTTTTATAT 16  
 |||||  
 Db 548 CCAAGCTTTTATAT 562

RESULT 15  
 US-09-949-016-130056  
 ; Sequence 130056, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C1001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 130056  
 ; LENGTH: 601  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-949-016-130056

Query Match 83.3%; Score 15; DB 4; Length 601;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTATAT 16  
 |||||  
 Db 548 CCAAGCTTTTATAT 562

Search completed: February 10, 2005, 08:50:04  
 Job time : 6.53136 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47.34 ; Search time 18.6055 Seconds  
(Without alignments)  
5727.097 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18

Sequence: 1 accaagctttatcataag 18

Scoring table: OLIGO-NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2002as:\*
- 6: geneseqn2002bs:\*
- 7: geneseqn2002as:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                 |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1          | 18    | 100.0       | 18     | 6     | ABK15249 Corn tran          |
| 2          | 18    | 100.0       | 1183   | 6     | ABK15245 Corn nk60          |
| 3          | 16    | 88.9        | 33     | 6     | ABL53728 Human BRG          |
| 4          | 16    | 88.9        | 107320 | 11    | ACN44736 Mouse gen          |
| 5          | 15    | 83.3        | 33     | 6     | ABV74652 Human zin          |
| 6          | 15    | 83.3        | 468    | 4     | AAH34926 Human col          |
| 7          | 15    | 83.3        | 549    | 6     | ABK76160 Bacillus           |
| 8          | 15    | 83.3        | 786    | 12    | ADK17054 Nanoarcha          |
| 9          | 15    | 83.3        | 1137   | 10    | ADK30185 Human nov          |
| 10         | 15    | 83.3        | 1248   | 12    | ADK30185 Human nov          |
| 11         | 15    | 83.3        | 2152   | 3     | AAA48446 Aa448446 Zebrafish |
| 12         | 15    | 83.3        | 2152   | 10    | ADC42306 Aa442306 CDNA enco |
| 13         | 15    | 83.3        | 2152   | 4     | ADH61060 Zebrafish          |
| 14         | 15    | 83.3        | 5248   | 4     | ABL10414 Drosophila         |
| 15         | 15    | 83.3        | 8897   | 4     | ABL20304 Drosophila         |
| 16         | 15    | 83.3        | 15932  | 8     | ABZ73902 Secreted           |
| 17         | 15    | 83.3        | 15932  | 8     | ADA44289 Human sec          |
| 18         | 15    | 83.3        | 90885  | 12    | ADK15049 Continuation (5 of |
| 19         | 15    | 83.3        | 144792 | 10    | ADC87620 Human GPC          |
| 20         | 15    | 83.3        | 218336 | 8     | ABQ76678 Androgen           |

|   |    |    |      |     |    |                    |
|---|----|----|------|-----|----|--------------------|
| C | 21 | 14 | 77.8 | 337 | 6  | ABN75127 Human ORF |
| C | 22 | 14 | 77.8 | 371 | 4  | AAK67516 Human imm |
| C | 23 | 14 | 77.8 | 377 | 4  | AAK67516 Human imm |
| C | 24 | 14 | 77.8 | 443 | 9  | ACH29826 Human tes |
| C | 25 | 14 | 77.8 | 453 | 4  | AAI11775 Probe #17 |
| C | 26 | 14 | 77.8 | 453 | 4  | ABA53465 Human foe |
| C | 27 | 14 | 77.8 | 453 | 4  | AAI33087 Probe #17 |
| C | 28 | 14 | 77.8 | 453 | 4  | ABA43053 Human bre |
| C | 29 | 14 | 77.8 | 453 | 4  | ABA23236 Probe #17 |
| C | 30 | 14 | 77.8 | 453 | 4  | AAK71192 Human bon |
| C | 31 | 14 | 77.8 | 453 | 4  | AAK01737 Human bra |
| C | 32 | 14 | 77.8 | 453 | 4  | ABS26774 Human liv |
| C | 33 | 14 | 77.8 | 453 | 5  | AAI01706 Human #16 |
| C | 34 | 14 | 77.8 | 453 | 6  | ABS01738 Human gen |
| C | 35 | 14 | 77.8 | 457 | 3  | AAK36726 Arabidops |
| C | 36 | 14 | 77.8 | 470 | 3  | AAK40800 Arabidops |
| C | 37 | 14 | 77.8 | 470 | 3  | AAK45230 Arabidops |
| C | 38 | 14 | 77.8 | 483 | 4  | AAI33833 Probe #25 |
| C | 39 | 14 | 77.8 | 483 | 6  | ABS02355 Human gen |
| C | 40 | 14 | 77.8 | 488 | 6  | ABS67082 Human col |
| C | 41 | 14 | 77.8 | 498 | 6  | ABT05259 Human col |
| C | 42 | 14 | 77.8 | 528 | 10 | ABX57529 Arabidops |
| C | 43 | 14 | 77.8 | 636 | 3  | AAK14827 Aspergill |
| C | 44 | 14 | 77.8 | 687 | 2  | AAV74919 Staphyloc |
| C | 45 | 14 | 77.8 | 761 | 6  | ABT09294 Phase-1 R |

#### ALIGNMENTS

|          |   |                       |
|----------|---|-----------------------|
| RESULT 1 | ABK15249  | standard; DNA; 18 BP. |
| ID       | ABK15249  |                       |
| XX       | ABK15249;   |                       |
| AC       | 29-AUG-2003 (revised)   |                       |
| DT       | 08-MAY-2002 (first entry)   |                       |
| XX       | Corn transgene junction probe #4.                                       |                       |
| DE       | Probe; ss; rice actin 1 promoter; RAI; RAI intron;                      |                       |
| KW       | chloroplast transit peptide gene; glycosylase resistance; RAI;          |                       |
| KW       | 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;      |                       |
| KW       | transcriptional terminator; cauliflower mosaic virus 35S promoter;      |                       |
| KW       | PV-ZMG732; transgenic; nk603.   |                       |
| XX       | Agrobacterium tumefaciens.  |                       |
| OS       | Oryza sativa.   |                       |
| OS       | Chimeric.   |                       |
| XX       | EP1167531-A1.   |                       |
| PN       | 02-JAN-2002.  |                       |
| PD       | 15-JUN-2001; 2001EP-00202314.   |                       |
| XX       | 22-JUN-2000; 2000US-0213567P.   |                       |
| PF       | 13-OCT-2000; 2000US-0240014P.   |                       |
| PR       | 13-OCT-2000; 2000US-0241215P.   |                       |
| PR       | (MONS ) MONSANTO TECHNOLOGY LLC.  |                       |
| PA       | Behr CF, Hironaka C, Heck GR, You J;                                    |                       |
| PI       | WPI, 2002-165871/22.  |                       |
| DR       | Novel DNA construct useful for producing a corn plant that tolerates    |                       |
| XX       | application of glyphosate herbicide, comprises two transgene expression |                       |
| PT       | cassettes.  |                       |
| PT       | Claim 8; Page 3; 25pp; English.   |                       |
| XX       |   |                       |

CC The invention relates to a DNA construct (I) comprising a first (F1) and  
CC a second (F2) expression cassette. F1 of the DNA construct, in operable  
CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
CC PV-ZMG32(nk603) of 498 or 1183bp, or primers and probes derived from  
CC them (used to detect transgene junction points). The construct is useful  
CC for producing a corn plant that tolerates application of glyphosate  
CC herbicide by transforming a corn cell with, selecting the corn cell for  
CC tolerance to application of glyphosate, and growing the corn cell into a  
CC fertile corn plant. The present sequence is a transgene junction probe  
CC sequence consisting of Agrobacterium and Rice sequences. (Updated on 29-  
CC AUG-2003 to standardise OS field)

XX  
SQ Sequence 18 BP; 7 A; 3 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 8.4;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTAATAAG 18  
DB 1 ACCAAGCTTTAATAAG 18

RESULT 2  
ABK15245  
ID ABK15245 standard; DNA; 1183 BP.  
XX  
AC ABK15245;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE Corn nk603 3' transgene/genomic sequence PCR fragment.  
XX  
KM ds; rice actin 1 promoter; RA1; RA1 intron;  
KM chloroplast transit peptide gene; glyphosate resistance; corn;  
KM 5-enol-pyruvylshikimate-3-phosphate synthase; EPSPS; Hsp70 intron;  
KM transcriptional terminator; cauliflower mosaic virus 35S promoter;  
KM PV-ZMG32; transgenic; nk603.  
XX  
XX Zea mays.  
OS Escherichia coli.  
OS Agrobacterium tumefaciens.  
OS Synthetic.  
OS Chimeric.  
XX  
XX Key Location/Qualifiers  
FH misc\_feature 1..164  
FT /\*tag= a  
FT /note= "T-AGRTU.nos vector sequence"  
FT 165..381  
FT /\*tag= b  
FT /note= "Vector sequence"  
FT 382..686  
FT /\*tag= c  
FT /note= "Corn plastid genes rps11/rpoA partial sequence"  
FT 687..1183  
FT /\*tag= d  
FT /note= "Corn genomic sequence"  
XX  
PN BP1167531-A1.  
XX  
PD 02-JAN-2002.  
XX  
PF 15-JUN-2001; 2001EP-00202314.  
XX  
PR 22-JUN-2000; 2000US-0213567P.  
PR 13-OCT-2000; 2000US-0240014P.  
PR 13-OCT-2000; 2000US-0241215P.  
XX

PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
XX Behr CF, Hironaka C, Heck GR, You J;  
XX WPI; 2002-165871/22.  
XX  
XX Novel DNA construct useful for producing a corn plant that tolerates  
PT application of glyphosate herbicide, comprises two transgene expression  
PT cassettes.  
XX  
PS Claim 3; Page 16-17; 25pp; English.

CC The invention relates to a DNA construct (I) comprising a first (F1) and  
CC a second (F2) expression cassette. F1 of the DNA construct, in operable  
CC linkage, has rice actin 1 (RA1) promoter, RA1 intron, chloroplast transit  
CC peptide gene (II), glyphosate tolerant 5-enol-pyruvylshikimate-3-  
CC phosphate synthase (EPSPS) gene (III), and transcriptional terminator  
CC (IV), and F2 has cauliflower mosaic virus 35S promoter, Hsp70 intron,  
CC (II), (III) and (IV). Also included are two DNA sequences from corn plant  
CC PV-ZMG32(nk603) of 498 or 1183bp, or primers and probes derived from  
CC them (used to detect transgene junction points). The construct is useful  
CC for producing a corn plant that tolerates application of glyphosate  
CC herbicide by transforming a corn cell with, selecting the corn cell for  
CC tolerance to application of glyphosate, and growing the corn cell into a  
CC fertile corn plant. The present sequence is the 1183bp amplicon  
CC consisting of the Agrobacterium transcriptional terminator sequence,  
CC vector sequences and corn sequences

XX  
SQ Sequence 1183 BP; 295 A; 289 C; 280 G; 319 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTAATAAG 18  
DB 156 ACCAAGCTTTAATAAG 173

RESULT 3  
ABL53728  
ID ABL53728 standard; DNA; 33 BP.  
XX  
AC ABL53728;  
XX  
DT 24-JUN-2002 (first entry)  
XX  
DE Human ERG conjugated protein 72 PCR primer #4.  
XX  
XX ERG conjugated protein; human; leukaemia; lipoma; inflammation;  
KM haemopathy; HIV infection; cytostatic; antiinflammatory; haemostatic;  
KM anti-HIV; virucide; gene therapy; PCR; primer; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX CN1329029-A.  
XX  
XX 02-JAN-2002.  
PD  
XX 19-JUN-2000; 2000CN-00116558.  
PF  
XX 19-JUN-2000; 2000CN-00116558.  
PR  
XX 19-JUN-2000; 2000CN-00116558.  
XX  
XX (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.  
XX  
XX Mao Y, Xie Y;  
XX  
XX WPI; 2002-305400/35.  
XX  
XX  
XX A novel polypeptide-human ERG conjugated protein 72 and polynucleotide  
PT for coding this polypeptide.  
XX  
XX Example 5; Page 18 (Disclosure); 33pp; Chinese.  
XX

XX The present invention relates to novel human ERG conjugated protein 72  
 CC (see ABB75713). The protein and its coding sequence are useful for  
 CC treatment of several diseases, such as medullary leukaemia, lipoma,  
 CC inflammation, immunological disease, haemopathy, development disturbance  
 CC and HIV infection. The present sequence is a PCR primer, which was used  
 CC in an example from the invention  
 XX

XX Sequence 33 BP; 8 A; 6 C; 1 G; 18 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 81; Mismatches 0; Indels 0; Gaps 0;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CCAAGCTTTATATA 17  
 DB 2 CCAAGCTTTATATA 17

RESULT 4  
 ACN44736/c  
 ID ACN44736 standard; DNA; 107320 BP.

AC ACN44736;

DT 18-NOV-2004 (first entry)

DE Mouse genomic sequence MCG2943.

KM Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

OS Mus musculus.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
 PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 1333; Ovp; English.

XX The present invention relates to novel DNA and protein sequences which  
 CC are associated with carcinomas. The sequences are useful for: (i) for  
 CC screening drug candidates; (ii) for screening of bioactive agent capable  
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biolchip;  
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
 CC carcinoma including lymphoma. The present sequence is one such CA coding  
 CC sequence. Note: This patent is an equivalent to basic patent  
 CC US2002182586A1, for which no sequence data was published  
 CC

XX Sequence 107320 BP; 30076 A; 19214 C; 21901 G; 35227 T; 0 U; 902 Other;

Query Match 88.9%; Score 16; DB 11; Length 107320;  
 Best Local Similarity 100.0%; Pred. No. 36; Mismatches 0; Indels 0; Gaps 0;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CCAAGCTTTATATA 17  
 DB 79987 CCAAGCTTTATATA 79972

RESULT 5  
 ABV74652  
 ID ABV74652 standard; DNA; 33 BP.

AC ABV74652;

DT 03-FEB-2003 (first entry)

DE Human zinc finger protein 64.90 PCR primer #4.

KM Human; zinc finger protein 64.90; cancer; HIV infection; cytostatic;  
 KM anti-HIV; PCR; primer; ss.

OS Homo sapiens.

EN CN1343696-A.

PD 10-APR-2002.

PF 19-SEP-2000; 2000CN-00125223.

PR 19-SEP-2000; 2000CN-00125223.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

PI Mao Y, Xie Y;

DR WPI; 2002-548869/59.

PT Novel human zinc finger protein 64.90 for treating cancer and HIV  
 PT infection.

PS Example 5; Page 18 (Disclosure); 34pp; Chinese.

XX The present invention relates to Human zinc finger protein 64.90 (see  
 CC ABB98772). The zinc finger protein and its coding sequence are useful for  
 CC treating cancer and HIV infection. The present sequence is a PCR primer,  
 CC which was used in an example from the invention  
 CC

XX Sequence 33 BP; 6 A; 9 C; 4 G; 14 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 6; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CCAAGCTTTATATA 16  
 DB 2 CCAAGCTTTATATA 16

RESULT 6  
 AAH34926  
 ID AAH34926 standard; cDNA; 468 BP.

AC AAH34926;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:2008.

KM Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KM colorectal carcinoma; chromosome 17; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.  
 PF 29-SEP-1999; 99US-0157137P.  
 PR 03-NOV-1999; 99US-0163280P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Ruben SM, Barash SC, Birse CE, Rosen CA;  
 PI WPI: 2001-235357/24.  
 DR P-PSDB; AAG75521.  
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 PT useful for preventing, diagnosing and/or treating colorectal cancers.  
 PS Claim 1; Page 3505; 9803pp; English.  
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
 CC proteins are collectively known as colon cancer antigens. The colon  
 CC cancer antigens have cytostatic activity and can be used in gene therapy  
 CC and vaccine production. N and P may be used in the prevention, diagnosis  
 CC and treatment of diseases associated with inappropriate P expression. For  
 CC example, N and P may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of P by expressing inactive proteins or to  
 CC supplement the patient's own production of P. Additionally, N may be used  
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acid  
 CC into a host cell and culturing the cell to express the proteins. N and P  
 CC can be used in the prevention, diagnosis and treatment of colorectal  
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
 CC sequences used in the exemplification of the present invention. N.B.  
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
 CC to 1052, 7921 and 7922  
 XX SQ Sequence 468 BP; 87 A; 140 C; 134 G; 99 T; 0 U; 8 Other;  
 XX  
 XX Query Match 83.3%; Score 15; DB 4; Length 468;  
 XX Best Local Similarity 100.0%; Pred. No. 2e+02;  
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAAGCTTTTATTA 15  
 Db 454 ACCAAGCTTTTATTA 468

RESULT 7  
 ABR76160/c  
 ID ABR76160 standard; DNA; 549 BP.  
 XX  
 AC ABR76160;  
 XX  
 DT 13-AUG-2002 (first entry)  
 XX  
 DE Bacillus licheniformis genomic sequence tag (GST) #3451.  
 XX  
 KW Differential gene expression; genomic sequenced tag; GST;  
 KW altered culture condition; environmental stress;  
 KW physiological provocation; ds.  
 XX  
 OS Bacillus licheniformis.  
 XX  
 PN WO200229113-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 05-OCT-2001; 2001WO-US031437.  
 XX  
 PR 06-OCT-2000; 2000US-00680598.  
 PR 27-MAR-2001; 2001US-0279526P.  
 XX

PA (NOVO) NOVOZYMES BIOTECH INC.  
 PA (NOVO) NOVOZYMES AS.  
 XX  
 PI Becka R, Clausen IG;  
 XX  
 DR WPI: 2002-416684/44.  
 XX  
 XX Monitoring differential expression of several genes in first Bacillus  
 PT cell relative to expression of same genes in one or more second Bacillus  
 PT cells, by using substrate containing Bacillus genomic sequenced tag  
 PT array.  
 XX  
 XX Claim 4; SEQ ID NO 3451; 200pp; English.  
 XX  
 XX The invention describes a method of monitoring differential expression of  
 CC genes in a first Bacillus cell relative to expression of the genes in  
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes  
 CC isolated from Bacillus cells to a substrate containing array of Bacillus  
 CC genomic sequenced tags (GST), examining the array, and determining  
 CC relative gene expression by an observed hybridisation reporter signal of  
 CC a spot in the array. The method is useful for measuring the expression of  
 CC genes in a first Bacillus cell relative to expression of the same genes  
 CC in one or more second Bacillus cells. The method is useful for monitoring  
 CC global expression of several genes from a Bacillus cell, discovering new  
 CC genes, identifying possible functions of unknown open reading frames and  
 CC monitoring gene copy number variation and stability. Monitoring changes  
 CC in expression of genes may be used to provide a representation of the way  
 CC in which Bacillus cells adapt to changes in culture conditions,  
 CC environmental stress or other physiological provocation. Extensive follow  
 CC up characterisation is unnecessary, when one spot on an array equals one  
 CC gene or one open reading frame, since sequence information is available.  
 CC This sequence represents a genomic sequence tag (GST) used in the method  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX SQ Sequence 549 BP; 145 A; 151 C; 118 G; 135 T; 0 U; 0 Other;  
 XX  
 XX Query Match 83.3%; Score 15; DB 6; Length 549;  
 XX Best Local Similarity 100.0%; Pred. No. 2e+02;  
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCAAGCTTTTATTAAT 16  
 Db 504 CCAAGCTTTTATTAAT 490

RESULT 8  
 ADK17054  
 ID ADK17054 standard; DNA; 786 BP.  
 XX  
 AC ADK17054;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Nanoarchaeum equitans cancer-associated (CA) gene #503.  
 XX  
 KW cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia;  
 KW ds; gene.  
 XX  
 OS Nanoarchaeum equitans.  
 XX  
 PN WO2003093434-A2.  
 XX  
 PD 13-NOV-2003.  
 XX  
 PF 01-MAY-2003; 2003WO-US013699.  
 XX  
 PR 01-MAY-2002; 2002US-0377447P.  
 XX  
 PA (DIVE-) DIVERSA CORP.  
 PA Stetter KO, Waters E, Kretz K, Podar M, Richardson T;  
 XX

PI Noordewier M;  
XX  
XX WPI; 2004-053041/05.  
DR P-PSDB; ADK17055.  
XX  
XX New recombinant cancer-associated genes, such as KCMN9, useful for  
PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,  
PT cervical, or skin cancers, lymphomas, or leukemia.  
XX  
XX Claim 5; SEQ ID NO 1006; 251pp; English.  
XX  
XX The invention comprises then amino acid and coding sequences of cancer-  
CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention  
CC also comprises the Nanoarchaeum equitans genome. The DNA and protein  
CC sequences of the invention are useful for diagnosing and treating cancer  
CC (e.g. carcinoma, lymphoma, or leukemia). The present DNA sequence  
CC represents a Nanoarchaeum equitans CA gene of the invention.  
XX  
XX Sequence 786 BP; 295 A; 112 C; 135 G; 244 T; 0 U; 0 Other;  
SQ  
Query Match 83.3%; Score 15; DB 12; Length 786;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 CAAGCTTTTATATA 17  
DB 584 CAAGCTTTTATATA 598  
RESULT 9  
AD30185/c  
ID AD30185 standard; cDNA; 1137 BP.  
XX  
XX ADC30185;  
AC  
XX 18-DEC-2003 (first entry)  
DT  
DE Human novel cDNA sequence, SEQ ID NO:267.  
XX  
XX Human; diagnostic; drug screening; forensics; gene mapping;  
KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KM ulcers; osteoporosis; autoimmune disease; cancer;  
KM molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KM neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;  
KM antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KM gene therapy; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO2003029271-A2.  
PN  
XX 10-APR-2003.  
PD  
XX 24-SEP-2002; 2002WO-US030474.  
PF  
XX 24-SEP-2001; 2001US-0324631P.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QH, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX  
XX WPI; 2003-371981/35.  
DR P-PSDB; ADC31156.  
XX  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
XX Claim 1; SEQ ID NO 267; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human cDNA sequence of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1137 BP; 295 A; 260 C; 292 G; 290 T; 0 U; 0 Other;  
Query Match 83.3%; Score 15; DB 10; Length 1137;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCAAGCTTTTATATA 15  
DB 398 ACCAAGCTTTTATATA 384  
RESULT 10  
AD035611/c  
ID AD035611 standard; DNA; 1248 BP.  
XX  
XX ADO35611;  
AC  
XX 26-AUG-2004 (first entry)  
DT  
DE Novel mouse gene sequence #284.  
XX  
XX mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;  
KM ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;  
KM viral disorder; ds; gene.  
XX  
XX Mus sp.  
OS  
XX WO2004046310-A2.  
PN  
XX 03-JUN-2004.  
PD  
XX 24-OCT-2003; 2003WO-US033948.  
PF  
XX 15-NOV-2002; 2002US-0426915P.  
PR 04-DEC-2002; 2002US-0431158P.  
PR 05-DEC-2002; 2002US-0431445P.  
PR 05-DEC-2002; 2002US-0431460P.  
PR 09-JUN-2003; 2003US-0476621P.  
PR 09-JUN-2003; 2003US-0476632P.

PR 08-JUL-2003; 2003US-0485217P.  
PR 08-JUL-2003; 2003US-0485359P.  
PR 08-AUG-2003; 2003US-0493332P.  
PR 08-AUG-2003; 2003US-0493356P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Heestir K, Hayashizaki Y, Kamiya M;  
XX WPI; 2004-431966/40.  
DR  
PT New mouse nucleic acid molecules and polypeptides, useful for treating  
XX cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart  
PT disease or thrombosis.  
XX  
PS Claim 1; SEQ ID NO 284; 263bp; English.  
XX  
CC The invention comprises 744 novel mouse DNA sequences (genes). The DNA  
CC sequences of the invention are useful for treating cancer, psoriasis,  
CC ulcerative colitis, inflammation, ischemic heart disease, thrombosis,  
CC immune disorder, bacterial disorders and viral disorders. The present  
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The  
CC present DNA sequence is not shown in the specification, but has been  
CC retrieved from the WIFO website.  
XX  
SQ Sequence 1248 BP; 452 A; 245 C; 236 G; 315 T; 0 U; 0 Other;  
XX  
Query Match 83.3%; Score 15; DB 12; Length 1248;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 CCAAGCTTTTATAT 16  
Db 931 CCAAGCTTTTATAT 917  
XX  
RESULT 11  
AAA48446/c  
ID AAA48446 standard; cDNA; 2152 BP.  
XX  
AC AAA48446;  
XX  
DT 15-SEP-2003 (revised)  
DT 27-OCT-2000 (first entry)  
XX  
DE Zebrafish parathyroid hormone type-3 receptor PTH3R coding sequence.  
XX  
KW Zebrafish; parathyroid hormone type-3 receptor; PTH3R;  
KW developmental disorder; physiological disorder; neurological disorder;  
KW ss.  
XX  
OS Danio rerio.  
XX  
FH Key Location/Qualifiers  
FT CDS 394..2022  
FT /\*tag= a  
FT /product= "PTH3R"  
XX  
PN MO200032775-A1.  
XX  
PD 08-JUN-2000.  
XX  
PF 30-NOV-1999; 99WO-US028207.  
XX  
PR 30-NOV-1998; 98US-0110467P.  
XX  
PA (JUEP/) JUEPPNER H.  
PA (RUBI/) RUBIN D A.  
XX  
PI Jueppner H, Rubin DA;  
XX WPI; 2000-412323/35.  
DR P-PSDB; AAY99601.  
DR

XX  
PT New nucleic acids encoding parathyroid hormone receptors PTH3R and PTH3R,  
PT useful for treating diseases or disorders associated with impaired  
PT receptor functions comprises a specific nucleotide sequence.  
XX  
PS Claim 23; Fig 1D; 111bp; English.  
XX  
CC The present sequence is the parathyroid hormone type-3 receptor (PTH3R)  
CC coding sequence from the zebrafish. It was obtained by sequencing a cDNA  
CC clone. The gene and protein can be used to detect diseases in man where  
CC the receptor is either overexpressed or underexpressed, and they can be  
CC used to treat these diseases, which may be developmentally, physiological  
CC or neurological disorders. They can also be used to identify agonists and  
CC antagonists which can be used in a similar manner. In addition, the gene  
CC can be used for chromosome identification. (Updated on 15-SEP-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 2152 BP; 592 A; 471 C; 529 G; 559 T; 0 U; 1 Other;  
XX  
Query Match 83.3%; Score 15; DB 3; Length 2152;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACCAAGCTTTTATTA 15  
Db 2090 ACCAAGCTTTTATTA 2076  
XX  
RESULT 12  
ADC42306/c  
ID ADC42306 standard; cDNA; 2152 BP.  
XX  
AC ADC42306;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE cDNA encoding zebrafish parathyroid hormone receptor PTH3R seq id 3.  
XX  
KW parathyroid hormone; PTH; PTH-related peptide; PTHrP;  
KW parathyroid hormone receptor; PTHR; chromosome identification; zebrafish;  
KW PTH3R; gene; ss.  
XX  
OS Danio rerio.  
XX  
FH Key Location/Qualifiers  
FT CDS 394..2022  
FT /\*tag= a  
FT /product= "Zebrafish PTH3R"  
FT /note= "Parathyroid hormone receptor"  
XX  
PN US6541220-B1.  
XX  
PD 01-APR-2003.  
XX  
PF 30-NOV-1999; 99US-00449632.  
XX  
PR 30-NOV-1998; 98US-0110467P.  
XX  
PA (GEO ) GEN HOSPITAL CORP.  
XX  
PI Jueppner H, Rubin DA;  
XX WPI; 2003-754511/71.  
DR P-PSDB; ADC42307.  
XX  
PT Novel nucleic acid comprising a polynucleotide encoding parathyroid  
PT hormone/parathyroid hormone-related peptide receptor, useful for  
PT chromosome identification.  
XX  
PS Disclosure; SEQ ID NO 3; 52bp; English.  
XX  
CC The invention describes an isolated nucleic acid (I) comprising a  
CC polynucleotide having a nucleotide sequence chosen from nucleotide



CC sequence encoding a parathyroid hormone (PTH)/PTH-related peptide (PTHrP)  
 CC receptor (PTHrR receptor) having a fully defined sequence of 536 amino  
 CC acids as given in the specification, PTHrR receptor, mature PTHrR  
 CC receptor, PTHrR extracellular or transmembrane domain, and their  
 CC complement. (I) is useful for diagnosing and treating decrease in the  
 CC standard or normal level of PTHrR receptor activity in an individual, and  
 CC for chromosome identification. This sequence encodes zebrafish PTHrR.  
 XX

SO Sequence 2152 BP; 592 A; 470 C; 529 G; 560 T; 0 U; 1 Other;  
 Query Match 83.3%; Score 15; DB 10; Length 2152;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAGCTTTATTA 15  
 |||||  
 DB 2090 ACCAGCTTTATTA 2076

RESULT 13  
 ADH61060/c  
 ID ADH61060 standard; cDNA; 2152 BP.

AC ADH61060;  
 DT 25-MAR-2004 (first entry)  
 XX

DE Zebrafish PTHrR cDNA from clone zPTHrR.

XX Zebrafish; parathyroid hormone; PTH; parathyroid hormone related protein;  
 KM PTHrP; parathyroid hormone related protein receptor; PTHr; PTHr;  
 KM diagnosis; prognosis; pharmaceutical composition; chromosome assay; gene;  
 ss.

OS Danio rerio.

XX Key location/Qualifiers  
 FH 394..2022  
 FT CDS

FT /\*tag= b  
 /product= "Zebrafish PTHrR protein"

FT sig\_peptide 394..457

FT /\*tag= a

FT mat\_peptide 458..2019

FT /\*tag= c  
 /product= "Zebrafish mature PTHrR protein"

FN US2003162256-A1.

XX 28-AUG-2003.

PD 25-FEB-2003; 2003US-00372095.

XX 30-NOV-1998; 98US-0110467P.

PR 30-NOV-1999; 99US-00449632.

XX (MARS-) MASSACHUSETTS GEN HOSPITAL.

PI Juepner H, Rubin DA;

XX WPI; 2003-897927/82.

DR P-PSDB; ADH61061.

PT New parathyroid hormone receptors designated PTHrR and PTHrR isolated  
 PT from zebrafish are useful to diagnose and treat parathyroid hormone  
 PT receptor-related diseases.

PS Claim 23; SEQ ID NO 3; 53pp; English.

XX The present invention relates to novel parathyroid hormone (PTH) and  
 CC parathyroid hormone related protein (PTHrP) receptors PTHrR and PTHrR  
 CC isolated from zebrafish. The invention is useful in the diagnosis and  
 CC prognosis of certain diseases and disorders that express significantly  
 CC decreased levels of PTHrR and PTHrR. The invention is also useful in

CC preparing pharmaceutical compositions and in chromosome assays. The  
 CC present sequence is zebrafish PTHrR cDNA.

SO Sequence 2152 BP; 592 A; 470 C; 529 G; 560 T; 0 U; 1 Other;

Query Match 83.3%; Score 15; DB 10; Length 2152;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAGCTTTATTA 15  
 |||||  
 DB 2090 ACCAGCTTTATTA 2076

RESULT 14  
 ABL10414/c  
 ID ABL10414 standard; cDNA; 5248 BP.

XX ABL10414;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 25724.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB66311.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX Claim 1; SEQ ID NO 25724; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutic and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL20511), expressed DNA  
 CC sequences (ABL101840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcr\_sequences

SO Sequence 5248 BP; 1525 A; 1056 C; 1162 G; 1505 T; 0 U; 0 Other;

Query Match 83.3%; Score 15; DB 4; Length 5248;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CAAGCTTTATTA 17  
 |||||  
 DB 4575 CAAGCTTTATTA 4561

```

RESULT 15
ID ABL20304/c
XX ABL20304 standard; DNA; 8897 BP.
AC ABL20304;
XX
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 12385.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 12385; 21pp + Sequence listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABL30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 8897 BP; 2656 A; 1738 C; 1855 G; 2648 T; 0 U; 0 Other;
SQ
Query Match 83.3%; Score 15; DB 4; Length 8897;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CAAGCTTTTATATA 17
DB 5922 CAAGCTTTTATATA 5908

```

Search completed: February 10, 2005, 13:44:15  
 Job time : 22.6055 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41 ; Search time 29.935 Seconds  
(without alignments)  
3460.943 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18

Sequence: 1 accaagctttacatag 18

Scoring table: OLIGO NUC  
Gapop\_60.0, Gapext 60.0

Searched: 4313806 seqs, 287787103 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09D\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description          |
|------------|-------|-------------|-----------|----|----------------------|
| 1          | 18    | 100.0       | 18        | 9  | US-09-872-051-12     |
| 2          | 18    | 100.0       | 18        | 18 | US-10-790-430-12     |
| 3          | 18    | 100.0       | 1183      | 9  | US-09-872-051-8      |
| 4          | 18    | 100.0       | 1183      | 18 | US-10-790-430-8      |
| 5          | 16    | 88.9        | 612       | 18 | US-10-425-115-76005  |
| 6          | 16    | 88.9        | 107320    | 13 | US-10-087-192-1333   |
| 7          | 15    | 83.3        | 440       | 17 | US-10-424-599-142571 |
| 8          | 15    | 83.3        | 468       | 15 | US-10-106-698-2018   |
| 9          | 15    | 83.3        | 533       | 13 | US-10-027-632-194059 |
| 10         | 15    | 83.3        | 533       | 17 | US-10-027-632-194059 |
| 11         | 15    | 83.3        | 549       | 9  | US-09-974-300-3451   |

|      |    |      |        |    |                      |                    |
|------|----|------|--------|----|----------------------|--------------------|
| C 12 | 15 | 83.3 | 613    | 13 | US-10-027-632-194858 | Sequence 194858,   |
| C 13 | 15 | 83.3 | 613    | 13 | US-10-027-632-194859 | Sequence 194859,   |
| C 14 | 15 | 83.3 | 613    | 17 | US-10-027-632-194858 | Sequence 194858,   |
| C 15 | 15 | 83.3 | 613    | 17 | US-10-027-632-194859 | Sequence 194859,   |
| C 16 | 15 | 83.3 | 996    | 13 | US-10-027-632-31469  | Sequence 31469, A  |
| C 17 | 15 | 83.3 | 996    | 17 | US-10-027-632-31469  | Sequence 31469, A  |
| C 18 | 15 | 83.3 | 2152   | 16 | US-10-372-095-3      | Sequence 3, Appl1  |
| C 19 | 15 | 83.3 | 9224   | 13 | US-10-108-605-254    | Sequence 254, Appl |
| C 20 | 15 | 83.3 | 94463  | 18 | US-10-719-993-6944   | Sequence 6944, Ap  |
| C 21 | 15 | 83.3 | 218336 | 18 | US-10-473-933-10     | Sequence 10, Appl  |
| C 22 | 15 | 83.3 | 561515 | 18 | US-10-741-601-5682   | Sequence 5682, Ap  |
| C 23 | 15 | 83.3 | 744802 | 17 | US-10-292-798-1369   | Sequence 1369, Ap  |
| C 24 | 14 | 77.8 | 194    | 18 | US-10-425-115-69101  | Sequence 69101, A  |
| C 25 | 14 | 77.8 | 201    | 18 | US-10-719-993-24777  | Sequence 24777, A  |
| C 26 | 14 | 77.8 | 201    | 18 | US-10-719-993-25079  | Sequence 25079, A  |
| C 27 | 14 | 77.8 | 201    | 18 | US-10-719-993-25081  | Sequence 25081, A  |
| C 28 | 14 | 77.8 | 206    | 18 | US-10-767-701-1656   | Sequence 1656, A   |
| C 29 | 14 | 77.8 | 290    | 18 | US-10-674-1248-12783 | Sequence 12783, A  |
| C 30 | 14 | 77.8 | 337    | 11 | US-09-864-408A-147   | Sequence 147, App  |
| C 31 | 14 | 77.8 | 430    | 17 | US-10-424-599-96924  | Sequence 96924, A  |
| C 32 | 14 | 77.8 | 431    | 18 | US-10-425-115-10849  | Sequence 10849, A  |
| C 33 | 14 | 77.8 | 432    | 18 | US-10-425-115-18546  | Sequence 18546, A  |
| C 34 | 14 | 77.8 | 443    | 10 | US-09-918-995-17038  | Sequence 17038, A  |
| C 35 | 14 | 77.8 | 453    | 9  | US-09-864-761-1702   | Sequence 1702, Ap  |
| C 36 | 14 | 77.8 | 469    | 17 | US-10-242-535A-39839 | Sequence 39839, A  |
| C 37 | 14 | 77.8 | 469    | 17 | US-10-085-783A-39839 | Sequence 39839, A  |
| C 38 | 14 | 77.8 | 482    | 13 | US-10-027-632-77058  | Sequence 77058, A  |
| C 39 | 14 | 77.8 | 482    | 17 | US-10-027-632-77058  | Sequence 77058, A  |
| C 40 | 14 | 77.8 | 483    | 9  | US-09-864-761-15448  | Sequence 15448, A  |
| C 41 | 14 | 77.8 | 489    | 17 | US-10-424-599-37844  | Sequence 37844, A  |
| C 42 | 14 | 77.8 | 498    | 9  | US-09-989-919-43     | Sequence 43, Appl  |
| C 43 | 14 | 77.8 | 528    | 9  | US-09-770-152-881    | Sequence 881, App  |
| C 44 | 14 | 77.8 | 611    | 13 | US-10-027-632-213307 | Sequence 213307,   |
| C 45 | 14 | 77.8 | 611    | 13 | US-10-027-632-213308 | Sequence 213308,   |

#### ALIGNMENTS

RESULT 1  
US-09-872-051-12  
Sequence 12, Application US/09872051  
Patent No. US20020013960A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Co  
APPLICANT: Behr, Carl  
APPLICANT: Hironaka, Catherine  
APPLICANT: Heck, Gregory  
APPLICANT: You, Jinong  
TITLE OF INVENTION: Corn Event PV-ZMGT32(Ink603) and Composition and Methods for Der  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 18  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(18)  
OTHER INFORMATION: Agrobacterium tumefaciens nos 3' terminator and rice actin prom  
US-09-872-051-12

Query Match 100.0%; Score 18; DB 9; Length 18;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18  
|||  
1 ACCAAGCTTTATATAG 18

## RESULT 2

US-10-790-430-12  
; Sequence 12, Application US/10790430  
; Publication No. US20040139493A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detect  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/10/790,430  
; PRIOR FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 18  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(18)  
; OTHER INFORMATION: Agrobacterium tumefaciens nos 3' terminator and rice actin promot  
US-10-790-430-12

Query Match 100.0%; Score 18; DB 18; Length 18;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18  
|||  
1 ACCAAGCTTTATATAG 18

## RESULT 3

US-09-872-051-8  
; Sequence 8, Application US/09872051  
; Patent No. US20020013960A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detect  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(1183)  
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
; OTHER INFORMATION: 165-381 construct vector DNA  
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA  
US-09-872-051-8

Query Match 100.0%; Score 18; DB 9; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18  
|||  
156 ACCAAGCTTTATATAG 173

## RESULT 4

US-10-790-430-8  
; Sequence 8, Application US/10790430  
; Publication No. US20040139493A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detect  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/10/790,430  
; PRIOR FILING DATE: 2004-03-01  
; PRIOR APPLICATION NUMBER: US/09/872,051  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(1183)  
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator  
; OTHER INFORMATION: 165-381 construct vector DNA  
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA  
US-10-790-430-8

Query Match 100.0%; Score 18; DB 18; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAAGCTTTATATAG 18  
|||  
156 ACCAAGCTTTATATAG 173

```
RESULT 5
US-10-425-115-76005/c
; Sequence 76005, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 76005
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169344C.1
US-10-425-115-76005

Query Match      88.9%; Score 16; DB 18; Length 612;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 CAAAGCTTTATATAG 18
DB      146 CAAAGCTTTATATAG 131

RESULT 6
US-10-087-192-1333/c
; Sequence 1333, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1333
; LENGTH: 107320
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (107320)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1333

Query Match      88.9%; Score 16; DB 13; Length 107320;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCAAGCTTTATATA 17
DB      79987 CCAAGCTTTATATA 79972

RESULT 7
US-10-424-599-142571/c
; Sequence 142571, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142571
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99757C.1
US-10-424-599-142571

Query Match      83.3%; Score 15; DB 17; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAGCTTTATATATAG 18
DB      264 AAGCTTTATATATAG 250

RESULT 8
US-10-106-698-2018
; Sequence 2018, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2018
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (434) - (434)
; OTHER INFORMATION: n equals a,t,c,g, or c
US-10-106-698-2018

Query Match      83.3%; Score 15; DB 15; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAAGCTTTATATA 15
DB      454 ACCAAGCTTTATATA 468

RESULT 9
US-10-027-632-194059
; Sequence 194059, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
```

```
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194059
LENGTH: 533
TYPE: DNA
ORGANISM: Human
US-10-027-632-194059
```

```
Query Match      83.3%; Score 15; DB 13; Length 533;
Best Local Similarity 100.0%; Pred. No. 1,7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
      |||||
DB      390 CCAAGCTTTTATAT 404
```

```
RESULT 10
US-10-027-632-194059
Sequence 194059, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194059
LENGTH: 533
TYPE: DNA
ORGANISM: Human
US-10-027-632-194059
```

```
Query Match      83.3%; Score 15; DB 17; Length 533;
Best Local Similarity 100.0%; Pred. No. 1,7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
      |||||
```

```
DB      390 CCAAGCTTTTATAT 404
```

```
RESULT 11
US-09-974-300-3451/C
Sequence 3451, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 10085,500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3451
LENGTH: 549
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-3451
```

```
Query Match      83.3%; Score 15; DB 9; Length 549;
Best Local Similarity 100.0%; Pred. No. 1,7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 CCAAGCTTTTATAT 16
      |||||
DB      504 CCAAGCTTTTATAT 490
```

```
RESULT 12
US-10-027-632-194858/C
Sequence 194858, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 194858
LENGTH: 613
TYPE: DNA
ORGANISM: Human
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(613)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-194858
```

Query Match 83.3% Score 15; DB 13; Length 613;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17  
|||  
Db 142 CAAGCTTTTATATA 128

RESULT 13  
US-10-027-632-194859/c

; Sequence 194859, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 194859  
; LENGTH: 613  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(613)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-194859

Query Match 83.3% Score 15; DB 13; Length 613;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17  
|||  
Db 142 CAAGCTTTTATATA 128

RESULT 14

US-10-027-632-194858/c  
; Sequence 194858, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 194858  
; LENGTH: 613  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(613)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-194858

Query Match 83.3% Score 15; DB 17; Length 613;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17  
|||  
Db 142 CAAGCTTTTATATA 128

RESULT 15

US-10-027-632-194859/c  
; Sequence 194859, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 194859  
; LENGTH: 613  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(613)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-194859

Query Match 83.3% Score 15; DB 17; Length 613;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CAAGCTTTTATATA 17  
|||  
Db 142 CAAGCTTTTATATA 128

Mon Feb 14 10:12:59 2005

us-10-790-430-12.oligo.rmpb

Page 6

Search completed: February 9, 2005, 11:12:47  
Job time : 30.935 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 121.68 Seconds  
(without alignments)  
5630.828 Million cell updates/sec

Title: US-10-790-430-12

Perfect score: 18  
Sequence: 1 accaagctttatcatag 18

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| C 1        | 17    | 94.4        | 651    | 9  | CE238229 tigr-gss- |
| C 2        | 16    | 88.9        | 118    | 5  | BP008680 BP008680  |
| C 3        | 16    | 88.9        | 196    | 5  | BM295435 BM295435  |
| C 4        | 16    | 88.9        | 199    | 1  | AV906194 AV906194  |
| C 5        | 16    | 88.9        | 280    | 9  | CI225702 CI225702  |
| C 6        | 16    | 88.9        | 298    | 1  | AV906393 AV906393  |
| C 7        | 16    | 88.9        | 357    | 8  | B39844 B39844      |
| C 8        | 16    | 88.9        | 363    | 1  | AV891743 AV891743  |
| C 9        | 16    | 88.9        | 382    | 8  | B55409 B55409      |
| C 10       | 16    | 88.9        | 431    | 8  | BA5104 BA5104      |
| C 11       | 16    | 88.9        | 432    | 8  | B98924 B98924      |
| C 12       | 16    | 88.9        | 476    | 8  | AQ829617 AQ829617  |
| C 13       | 16    | 88.9        | 527    | 1  | AV903514 AV903514  |
| C 14       | 16    | 88.9        | 572    | 1  | AV904180 AV904180  |
| C 15       | 16    | 88.9        | 618    | 1  | BP016541 BP016541  |
| C 16       | 16    | 88.9        | 627    | 5  | BP004638 BP004638  |
| C 17       | 16    | 88.9        | 635    | 1  | AV895582 AV895582  |
| C 18       | 16    | 88.9        | 640    | 8  | B20453 B20453      |
| C 19       | 16    | 88.9        | 651    | 5  | BM167398 BM167398  |
| C 20       | 16    | 88.9        | 655    | 1  | AV897075 AV897075  |
| C 21       | 16    | 88.9        | 657    | 5  | BM065072 BM065072  |
| C 22       | 16    | 88.9        | 659    | 5  | BM192868 BM192868  |
| C 23       | 16    | 88.9        | 665    | 5  | BM065061 BM065061  |
| C 24       | 16    | 88.9        | 666    | 1  | AV906223 AV906223  |

|      |    |      |      |   |                   |
|------|----|------|------|---|-------------------|
| C 25 | 16 | 88.9 | 667  | 5 | BM066300 BM066300 |
| C 26 | 16 | 88.9 | 668  | 1 | AV893354 AV893354 |
| C 27 | 16 | 88.9 | 668  | 5 | BM191586 BM191586 |
| C 28 | 16 | 88.9 | 669  | 1 | AV893491 AV893491 |
| C 29 | 16 | 88.9 | 670  | 5 | BP016388 BP016388 |
| C 30 | 16 | 88.9 | 672  | 5 | BP01597 BP01597   |
| C 31 | 16 | 88.9 | 674  | 5 | BP002271 BP002271 |
| C 32 | 16 | 88.9 | 678  | 5 | BP008293 BP008293 |
| C 33 | 16 | 88.9 | 682  | 5 | BP006864 BP006864 |
| C 34 | 16 | 88.9 | 683  | 5 | BM296842 BM296842 |
| C 35 | 16 | 88.9 | 695  | 5 | BP004801 BP004801 |
| C 36 | 16 | 88.9 | 717  | 6 | CA174264 CA174264 |
| C 37 | 16 | 88.9 | 834  | 8 | BZ466621 BZ466621 |
| C 38 | 16 | 88.9 | 862  | 5 | BU459296 BU459296 |
| C 39 | 16 | 88.9 | 928  | 6 | CA243532 CA243532 |
| C 40 | 16 | 88.9 | 1358 | 9 | AG463618 AG463618 |
| C 41 | 15 | 83.3 | 128  | 8 | CC380153 CC380153 |
| C 42 | 15 | 83.3 | 158  | 8 | B26837 B26837     |
| C 43 | 15 | 83.3 | 171  | 9 | AL943005 AL943005 |
| C 44 | 15 | 83.3 | 187  | 7 | Z38858 Z38858     |
| C 45 | 15 | 83.3 | 193  | 4 | BI490751 BI490751 |

#### ALIGNMENTS

RESULT 1  
CE238229/c 651 bp DNA linear GSS 25-SEP-2003  
LOCUS tigr-gss-dog-1700033406145 Dog Library Canis familiaris genomic,  
DEFINITION genomic survey sequence.

ACCESSION CE238229 GI:35393987  
VERSION CE238229.1  
KEYWORDS GSS.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris

REFERENCE 1 (bases 1 to 651)  
Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.  
The dog genome: survey sequencing and comparative analysis

TITLE JOURNAL Science 301 (5641), 1898-1903 (2003)  
MEDLINE 22875432  
PUBMED 14512627

COMMENT Contact: Kirkness EF

The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirkness@tigr.org  
Class: Shotgun.

FEATURES  
source location/Qualifiers

1..651  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: Bactx; Libraries were prepared from  
peripheral blood"

#### ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 651;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCAAGCTTTTAAATAG 18  
DB 606 CCAAGCTTTTAAATAG 590

RESULT 2  
BP008680 118 bp mRNA linear EST 15-MAR-2002  
LOCUS BP008680  
DEFINITION intestinalis cDNA clone ciad49h02 5', mRNA sequence.  
ACCESSION BP008680  
VERSION BP008680.1 GI:19500157  
KEYWORDS EST.  
SOURCE Ciona intestinalis  
ORGANISM Ciona intestinalis  
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
Phlebobranchia; Clonidae; Ciona.  
REFERENCE 1 (bases 1 to 118)  
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.  
TITLE Expressed genes in Ciona intestinalis  
JOURNAL Unpublished (2000)  
COMMENT Contact: Nori Satoh  
Department of Zoology  
Kyoto University  
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
Tel: 81-75-753-4081  
Fax: 81-75-705-1113  
Email: satoh@ascidian.zool.kyoto-u.ac.jp.  
location/Qualifiers  
1. 118  
/organism="Ciona intestinalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:7719"  
/clone="ciad49h02"  
/issue\_type="whole animal"  
/dev\_stage="young adult"  
/clone\_1lb="Nori Satoh unpublished cDNA library, young adult"

ORIGIN  
Query Match 88.9%; Score 16; DB 5; Length 118;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATAT 16  
|||||  
DB 21 ACCAAGCTTTATAT 36

RESULT 3  
BM295435 196 bp mRNA linear EST 11-NOV-2002  
LOCUS BM295435  
DEFINITION intestinalis cDNA clone cinc010p13 5', mRNA sequence.  
ACCESSION BM295435  
VERSION BM295435.1 GI:24876046  
KEYWORDS EST.  
SOURCE Ciona intestinalis  
ORGANISM Ciona intestinalis  
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
Phlebobranchia; Clonidae; Ciona.  
REFERENCE 1 (bases 1 to 196)  
AUTHORS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.  
TITLE Expressed genes in Ciona intestinalis (2002c)  
JOURNAL Unpublished (2002)  
COMMENT Contact: Nori Satoh  
Department of Zoology  
Kyoto University  
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
Tel: 81-75-753-4081  
Fax: 81-75-705-1113  
Email: satoh@ascidian.zool.kyoto-u.ac.jp.  
location/Qualifiers  
1. 196  
/organism="Ciona intestinalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:7719"

FEATURES  
source

ORIGIN  
Query Match 88.9%; Score 16; DB 5; Length 196;  
Best Local Similarity 100.0%; Pred. No. 16+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATAT 16  
|||||  
DB 54 ACCAAGCTTTATAT 39

RESULT 5  
CL225702 280 bp DNA linear GSS 08-JAN-2004  
LOCUS ZMMB0564122r ZMMB0564122r ZMMB0564122 3',  
DEFINITION genomic survey sequence.  
ACCESSION CL225702  
VERSION CL225702.1 GI:40761719  
KEYWORDS GSS.  
SOURCE Zee mays  
ORGANISM Zee mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zee.  
REFERENCE 1 (bases 1 to 280)  
/clone="cinc010p13"  
/issue\_type="neural complex"  
/clone\_1lb="Nori Satoh unpublished cDNA library, neural complex"

ORIGIN  
Query Match 88.9%; Score 16; DB 5; Length 199;  
Best Local Similarity 100.0%; Pred. No. 16+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAAGCTTTATAT 16  
|||||  
DB 54 ACCAAGCTTTATAT 39

RESULT 4  
AV906194/c 199 bp mRNA linear EST 09-NOV-2001  
LOCUS AV906194/c  
DEFINITION intestinalis cDNA clone rciad65c15 3', mRNA sequence.  
ACCESSION AV906194  
VERSION AV906194.1 GI:16895292  
KEYWORDS EST.  
SOURCE Ciona intestinalis  
ORGANISM Ciona intestinalis  
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
Phlebobranchia; Clonidae; Ciona.  
REFERENCE 1 (bases 1 to 199)  
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.  
TITLE Expressed genes in Ciona intestinalis  
JOURNAL Unpublished (2000)  
COMMENT Contact: Nori Satoh  
Department of Zoology  
Kyoto University  
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
Tel: 81-75-753-4081  
Fax: 81-75-705-1113  
Email: satoh@ascidian.zool.kyoto-u.ac.jp.  
location/Qualifiers  
1. 199  
/organism="Ciona intestinalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:7719"  
/clone="rciad65c15"  
/issue\_type="whole animal"  
/dev\_stage="young adult"  
/clone\_1lb="Nori Satoh unpublished cDNA library, young adult"

FEATURES  
source

**AUTHORS** Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,  
 Zohovetz, V., Ruke, G., Yu, Y., Wang, R. and Messing, J.  
**TITLE** Sequencing of the maize genome at FGR (2003c)  
**JOURNAL** Unpublished (2003)  
**COMMENT** Dr. Joachim Messing's lab  
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers  
 University  
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
 Tel: 732 445 3801  
 Fax: 732 445 5735  
 Email: bharti@waksman.rutgers.edu  
 Seq primer: SP6  
 Class: BAC ends.

**FEATURES**  
 source  
 1..280  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /culti\_vars="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMWB056412"  
 /lab\_host="E. coli DH10B"  
 /clone\_1lb="ZMWB056412"  
 /note="Vector: pTARBA1.3; Site\_1: BamHI; Site\_2: BamHI"

**ORIGIN**

Query Match 88.9%; Score 16; DB 9; Length 280;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 2 CCAAGCTTTTATATA 17  
 |||||  
 |||||  
**Db** 51 CCAAGCTTTTATATA 66

**RESULT 6**  
**LOCUS** AV906393/c 298 bp mRNA linear EST 09-NOV-2001  
**DEFINITION** AV906393 Nori Satoh unpublished cDNA library, young adult Ciona  
 intestinalis cDNA clone rcia465n15 3', mRNA sequence.  
**ACCESSION** AV906393  
**VERSION** AV906393.1 GI:16895491  
**KEYWORDS** EST.  
**SOURCE** Ciona intestinalis  
**ORGANISM** Ciona intestinalis  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 Plebobranchia; Cionidae; Ciona.

**REFERENCE** 1 (bases 1 to 298)  
**AUTHORS** Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.  
**TITLE** Expressed genes in Ciona intestinalis  
**JOURNAL** Unpublished (2000)  
**COMMENT** Contact: Nori Satoh  
 Department of Zoology  
 Kyoto University  
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4081  
 Fax: 81-75-705-1113  
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

**FEATURES**  
 source  
 1..298  
 /organism="Ciona intestinalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7719"  
 /clone="rcia465n15"  
 /tissue\_type="whole animal"  
 /dev\_stage="young adult"  
 /clone\_1lb="Nori Satoh unpublished cDNA library, young  
 adult"

**ORIGIN**

Query Match 88.9%; Score 16; DB 1; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 ACCAAGCTTTTATAT 16  
 |||||  
 |||||  
**Db** 80 ACCAAGCTTTTATAT 65

**RESULT 7**  
**B39844** 357 bp DNA linear GSS 18-OCT-1997  
**LOCUS** B39844  
**DEFINITION** HS-1050-B1-D08-MR.abi CIT Human Genomic Sperm Library C Homo  
 sapiens genomic clone plate=CT 772 Col=15 Row=H, genomic survey  
 sequence.  
**ACCESSION** B39844  
**VERSION** B39844.1 GI:2544096  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**REFERENCE** 1 (bases 1 to 357)  
**AUTHORS** Mahairas, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schmidt, S.,  
 Tralcoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.E.  
**TITLE** Construction of a Characterized Clone Resource for Genomic  
 Sequencing: Generation and Preliminary Analysis of 20,000 Sequence  
 Tagged Connectors  
 Unpublished (1997)  
**JOURNAL** Contact: Mahairas GG, Zackrone KD, Hood L  
 University of Washington  
 Seattle, WA 98195, USA  
 Tel: (206) 616-8744  
 Fax: (206) 685-7301  
 Email: kzackron@u.washington.edu  
 Sequence Tagged Connector  
 Plate: CT 772 row: H column: 15  
 Class: BAC ends  
 High quality sequence stop: 357.

**FEATURES**  
 source  
 1..357  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=CT 772 Col=15 Row=H"  
 /sex="M"  
 /clone\_1lb="CIT Human Genomic Sperm Library C"  
 /note="Organ: sperm; Vector: pBeloBAC11; BAC clones in  
 E-Coli DH10B"

**ORIGIN**

Query Match 88.9%; Score 16; DB 8; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3 CAAGCTTTTATATAG 18  
 |||||  
 |||||  
**Db** 138 CAAGCTTTTATATAG 153

**RESULT 8**  
**AV891743/c** 363 bp mRNA linear EST 09-NOV-2001  
**LOCUS** AV891743  
**DEFINITION** AV891743 Nori Satoh unpublished cDNA library, young adult Ciona  
 intestinalis cDNA clone rcia26f24 3', mRNA sequence.  
**ACCESSION** AV891743  
**VERSION** AV891743.1 GI:16880769  
**KEYWORDS** EST.  
**SOURCE** Ciona intestinalis  
**ORGANISM** Ciona intestinalis  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 Plebobranchia; Cionidae; Ciona.

**REFERENCE** 1 (bases 1 to 363)  
**AUTHORS** Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.  
**TITLE** Expressed genes in Ciona intestinalis  
**JOURNAL** Unpublished (2000)

## COMMENT

Contact: Nori Satoh  
Department of Zoology  
Kyoto University  
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
Tel: 81-75-753-4081  
Fax: 81-75-705-1113  
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

## FEATURES

## source

1.363  
/organism="Clona intestinalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:7719"  
/clone="rciad26124"  
/tissue\_type="whole animal"  
/dev\_stage="young adult"  
/clone\_lib="Nori Satoh unpublished cDNA library, young adult"

## ORIGIN

Query Match 88.9%; Score 16; DB 1; Length 363;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGCTTTATAT 16  
|||||  
DB 53 ACCAGCTTTATAT 38

## RESULT 9

B55409 382 bp DNA linear GSS 20-JUN-1998  
LOCUS CIT-HSP-387A19.TF CIT-HSP Homo sapiens genomic clone 387A19,  
DEFINITION genomic survey sequence.  
ACCESSION B55409  
VERSION B55409.1 GI:2609743  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 382)  
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linhar,K.,  
Golden,K., Berry,K., Granger,D., Sub,E., Whible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building  
Unpublished (1997)  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tcdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21  
Clase: BAC ends.

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1.382  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:5379336"  
/db\_xref="taxon:9606"  
/clone="387A19"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

## ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 382;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CAAGCTTTATATAG 18  
|||||  
DB 1 CAAGCTTTATATAG 16

## RESULT 10

B45104 431 bp DNA linear GSS 21-OCT-1997  
LOCUS HS-1060-B1-F03-NR.abi CIT Human Genomic Sperm Library C Homo  
DEFINITION sapiens genomic clone Plate=CT 782 Col=5 Row=L, genomic survey  
sequence.  
ACCESSION B45104  
VERSION B45104.1 GI:2549938  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 431)  
Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,  
Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.  
Construction of a Characterized Clone Resource for Genomic  
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence  
Tagged Connectors  
Unpublished (1997)  
Contact: Mahairas GG, Zackrone KD, Hood L  
University of Washington  
Seattle, WA 98195, USA  
Tel: (206) 616-8744  
Fax: (206) 685-7301  
Email: kzackron@u.washington.edu  
Sequence Tagged Connector  
Plate: CT 782 row: L column: 5  
Class: BAC ends  
High quality sequence stop: 431.  
Location/Qualifiers

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1.431  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=CT 782 Col=5 Row=L"  
/sex="M"  
/clone\_lib="CIT Human Genomic Sperm Library C"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in  
E-Coli DH10B"

## ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CAAGCTTTATATAG 18  
|||||  
DB 50 CAAGCTTTATATAG 65

## RESULT 11

B98924 432 bp DNA linear GSS 26-JUN-1998  
LOCUS CIT-HSP-2174P18.TF CIT-HSP Homo sapiens genomic clone 2174P18,  
DEFINITION genomic survey sequence.  
ACCESSION B98924  
VERSION B98924.1 GI:3026734  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 432)

## REFERENCE

**AUTHORS** Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and Venter, J.C.  
**TITLE** Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)  
**JOURNAL** Unpublished (1998)  
**COMMENT** Other\_GSSs: CIT-HSP-2174P18.TR  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
[http://www.tigr.org/tcdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html)  
 Seq primer: M13-21;  
 Class: BAC ends.

**FEATURES**  
 source  
 1. .432  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7104820"  
 /db\_xref="taxon:9606"  
 /clone="2174P18"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /clone\_lib="CIT-HSP"  
 /note="Vector: pBel0BAC11; Site\_1: HindIII; Site\_2: HindIII"

**ORIGIN**  
 Query Match 88.9%; Score 16; DB 8; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 3 CAAGCTTTTATATAG 18  
 |||||  
 |||||  
**Db** 29 CAAGCTTTTATATAG 44

**RESULT 12**  
 A0829617 476 bp DNA linear GSS 27-AUG-1999  
 A0829617 HS 4832\_A1 B01 SP68 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=4832 Col=1 Row=1, genomic survey sequence.  
**ACCESSION** A0829617 GI:5795679  
**VERSION** A0829617  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE** 1 (bases 1 to 476)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.  
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
**JOURNAL** 99380589  
**COMMENT** 10449764  
 Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).  
 BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 4832 row: 1 column: 1  
 Seq primer: r7  
 Class: BAC ends  
 High quality sequence stop: 476.  
**FEATURES**  
 source  
 1. .476  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=4832 Col=1 Row=1"  
 /sex="male"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: pBel0BAC11; BAC Clones in E-Coli DH10B"

**ORIGIN**  
 Query Match 88.9%; Score 16; DB 8; Length 476;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 3 CAAGCTTTTATATAG 18  
 |||||  
 |||||  
**Db** 177 CAAGCTTTTATATAG 192

**RESULT 13**  
 AV903514/c 527 bp mRNA linear EST 09-NOV-2001  
 AV903514/c LOCUS  
 DEFINITION AV903514 Nori Satoh unpublished cDNA library, young adult Ciona intestinalis cDNA clone rciad60k21 3', mRNA sequence.  
 AV903514  
**ACCESSION** AV903514 GI:16892612  
**VERSION** EST.  
**KEYWORDS** Ciona intestinalis  
**SOURCE** Ciona intestinalis  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Pleurobranchia; Cionidae; Ciona.  
**REFERENCE** 1 (bases 1 to 527)  
 Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.  
 Expressed genes in Ciona intestinalis  
 Unpublished (2000)  
**JOURNAL** Contact: Nori Satoh  
 Department of Zoology  
 Kyoto University  
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4081  
 Fax: 81-75-705-1113  
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

**FEATURES**  
 source  
 1. .527  
 /organism="Ciona intestinalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7719"  
 /clone="rciad60k21"  
 /issue\_type="whole animal"  
 /dev\_stage="young adult"  
 /clone\_lib="Nori Satoh unpublished cDNA library, young adult"

**ORIGIN**  
 Query Match 88.9%; Score 16; DB 1; Length 527;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 1 ACCAAGCTTTATAT 16  
 |||||  
 |||||  
**Db** 80 ACCAAGCTTTATAT 65

**RESULT 14**  
 AV904180 572 bp mRNA linear EST 09-NOV-2001  
 AV904180/c LOCUS  
 DEFINITION AV904180 Nori Satoh unpublished cDNA library, young adult Ciona

```

ACCESSION      Intestinalis cDNA clone rc1ad62115 3', mRNA sequence.
VERSION        AV904180
KEYWORDS       EST.
SOURCE         Ciona intestinalis
ORGANISM       Ciona intestinalis
REFERENCE      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
AUTHORS       Plebobranchia; Clonidae; Ciona.
TITLE          1 (bases 1 to 572)
JOURNAL        Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
COMMENT        Expressed genes in Ciona intestinalis
                Unpublished (2000)
                Contact: Nori Satoh
                Department of Zoology
                Kyoto University
                Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
                Tel: 81-75-753-4081
                Fax: 81-75-705-1113
                Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source         1..572
                /organism="Ciona intestinalis"
                /mol_type="mRNA"
                /db_xref="taxon:7719"
                /clone="rc1ad62115"
                /issue_type="whole animal"
                /dev_stage="young adult"
                /clone_lib="Nori Satoh unpublished cDNA library, young
                adult"

ORIGIN
Query Match      88.9%; Score 16; DB 1; Length 572;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAAGCTTTATAT 16
        |||||||
        42 ACCAAGCTTTATAT 27

RESULT 15
BP016541      618 bp mRNA linear EST 15-MAR-2002
LOCUS         BP016541 Nori Satoh unpublished cDNA library, young adult Ciona
DEFINITION    Intestinalis cDNA clone cladsn15 5', mRNA sequence.
ACCESSION     BP016541
VERSION       BP016541.1 GI:19508018
KEYWORDS      EST.
SOURCE        Ciona intestinalis
ORGANISM      Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Plebobranchia; Clonidae; Ciona.
1 (bases 1 to 618)
Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
Expressed genes in Ciona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source         1..618
                /organism="Ciona intestinalis"
                /mol_type="mRNA"
                /db_xref="taxon:7719"
                /clone="cladsn15"
                /issue_type="whole animal"
                /dev_stage="young adult"
                /clone_lib="Nori Satoh unpublished cDNA library, young
                adult"

```

```

ORIGIN
Query Match      88.9%; Score 16; DB 5; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAAGCTTTATAT 16
        |||||||
        592 ACCAAGCTTTATAT 607

Search completed: February 10, 2005, 17:02:02
Job time : 124.68 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:23:41, Search time 828.202 Seconds  
(without alignments)  
3460.943 Million cell updates/sec

Title: US-10-790-430-7

Perfect score: 498

Sequence: 1 aatcgatcccaaatcgcgcac.....aaaagctgcccgaagtga 498

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 4313806 seqs, 2877871033 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                   | Description        |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1          | 498   | 100.0       | 498    | 9 US-09-872-051-7    | Sequence 7, Appl1  |
| 2          | 498   | 100.0       | 498    | 18 US-10-790-430-7   | Sequence 7, Appl1  |
| 3          | 118   | 23.7        | 1183   | 9 US-09-872-051-8    | Sequence 8, Appl1  |
| 4          | 118   | 23.7        | 1183   | 18 US-10-790-430-8   | Sequence 8, Appl1  |
| 5          | 101   | 20.3        | 2378   | 15 US-10-213-791-27  | Sequence 27, Appl1 |
| 6          | 70    | 14.1        | 1501   | 14 US-10-012-070A-49 | Sequence 49, Appl1 |
| 7          | 70    | 14.1        | 1501   | 14 US-10-012-013-44  | Sequence 44, Appl1 |
| 8          | 68    | 13.7        | 1259   | 14 US-09-991-209-43  | Sequence 43, Appl1 |
| 9          | 68    | 13.7        | 2480   | 19 US-10-678-588A-1  | Sequence 1, Appl1  |
| 10         | 68    | 13.7        | 5164   | 10 US-09-991-209-36  | Sequence 36, Appl1 |
| 11         | 68    | 13.7        | 5277   | 10 US-09-991-209-25  | Sequence 25, Appl1 |

|    |    |      |       |                       |                      |
|----|----|------|-------|-----------------------|----------------------|
| 12 | 68 | 13.7 | 5295  | 10 US-09-991-209-38   | Sequence 38, Appl1   |
| 13 | 68 | 13.7 | 5327  | 10 US-09-991-209-27   | Sequence 27, Appl1   |
| 14 | 68 | 13.7 | 5337  | 10 US-09-991-209-19   | Sequence 19, Appl1   |
| 15 | 68 | 13.7 | 5337  | 10 US-09-991-209-23   | Sequence 23, Appl1   |
| 16 | 68 | 13.7 | 5337  | 10 US-09-991-209-31   | Sequence 31, Appl1   |
| 17 | 68 | 13.7 | 5338  | 10 US-09-991-209-15   | Sequence 15, Appl1   |
| 18 | 68 | 13.7 | 5338  | 10 US-09-991-209-17   | Sequence 29, Appl1   |
| 19 | 68 | 13.7 | 5345  | 10 US-09-991-209-41   | Sequence 41, Appl1   |
| 20 | 68 | 13.7 | 5387  | 10 US-09-991-209-21   | Sequence 21, Appl1   |
| 21 | 68 | 13.7 | 5395  | 10 US-09-991-209-21   | Sequence 21, Appl1   |
| 22 | 67 | 13.5 | 4032  | 9 US-09-970-921-5     | Sequence 5, Appl1    |
| 23 | 63 | 12.7 | 399   | 9 US-09-682-597-5     | Sequence 5, Appl1    |
| 24 | 63 | 12.7 | 399   | 18 US-10-727-423-5    | Sequence 5, Appl1    |
| 25 | 63 | 12.7 | 399   | 18 US-10-737-487-5    | Sequence 5, Appl1    |
| 26 | 62 | 12.4 | 6865  | 10 US-09-845-064-13   | Sequence 13, Appl1   |
| 27 | 62 | 12.4 | 10003 | 10 US-09-845-064-21   | Sequence 21, Appl1   |
| 28 | 60 | 6.0  | 30    | 9 US-09-682-597-2     | Sequence 2, Appl1    |
| 29 | 60 | 6.0  | 30    | 18 US-10-727-423-2    | Sequence 2, Appl1    |
| 30 | 60 | 6.0  | 30    | 18 US-10-737-487-2    | Sequence 2, Appl1    |
| 31 | 29 | 5.8  | 29    | 9 US-09-872-051-4     | Sequence 4, Appl1    |
| 32 | 29 | 5.8  | 29    | 18 US-10-790-430-4    | Sequence 4, Appl1    |
| 33 | 22 | 4.4  | 22    | 9 US-09-872-051-13    | Sequence 13, Appl1   |
| 34 | 22 | 4.4  | 22    | 9 US-09-872-051-14    | Sequence 14, Appl1   |
| 35 | 22 | 4.4  | 22    | 18 US-10-790-430-13   | Sequence 13, Appl1   |
| 36 | 22 | 4.4  | 22    | 18 US-10-790-430-14   | Sequence 14, Appl1   |
| 37 | 21 | 4.2  | 2000  | 17 US-10-260-238-2562 | Sequence 2562, Appl1 |
| 38 | 20 | 4.0  | 2107  | 15 US-10-213-791-29   | Sequence 29, Appl1   |
| 39 | 20 | 4.0  | 2436  | 15 US-10-213-791-31   | Sequence 31, Appl1   |
| 40 | 20 | 4.0  | 3039  | 15 US-10-232-665-19   | Sequence 19, Appl1   |
| 41 | 20 | 4.0  | 3039  | 15 US-10-232-665-21   | Sequence 21, Appl1   |
| 42 | 20 | 4.0  | 3044  | 15 US-10-232-665-17   | Sequence 17, Appl1   |
| 43 | 20 | 4.0  | 3455  | 15 US-10-232-665-18   | Sequence 18, Appl1   |
| 44 | 20 | 4.0  | 3455  | 15 US-10-232-665-16   | Sequence 16, Appl1   |
| 45 | 20 | 4.0  | 3469  | 15 US-10-232-665-23   | Sequence 23, Appl1   |

#### ALIGNMENTS

RESULT 1  
US-09-872-051-7  
; Sequence 7, Appl1 Application US/09872051  
; Patent No. US20020013960A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMGR32 (nke603) and Composition and Methods for Det  
; FILE REFERENCE: 38-21 (52258)B  
; CURRENT APPLICATION NUMBER: US/09/872,051  
; PRIOR APPLICATION NUMBER: 2001-06-01  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; NAME/KEY: source  
; LOCATION: (1)..(498)  
; OTHER INFORMATION: 1-304 Zea maize genomic DNA  
; OTHER INFORMATION: 305-349 construct vector DNA  
; OTHER INFORMATION: 350-498 rice actin 1 promoter DNA  
US-09-872-051-7

```

? ORIENTATION: MULTICUT Sequence
?
? FEATURE:
? NAME/KEY: SOURCE

```

LENGTH: 1183  
TYPE: DNA



```

; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
; US-09-872-051-8

Query Match      23.7%; Score 118; DB 9; Length 1183;
Best Local Similarity 100.0%; Pred. No. 9.6e-54;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      308 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGCTTAACAAGCTTACTCGAGGTCA 367
      381 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGCTTAACAAGCTTACTCGAGGTCA 322
Db      368 TTCATATGCTTGAGAAAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 425
      321 TTCATATGCTTGAGAAAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 264

RESULT 4
US-10-790-430-8/c
; Sequence 8, Application US/10790430
; Publication No. US20040139493A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMGT32(nk603) and Composition and Methods for Detecting and Compositions
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/10/790,430
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
; US-10-790-430-8

Query Match      23.7%; Score 118; DB 18; Length 1183;
Best Local Similarity 100.0%; Pred. No. 9.6e-54;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      308 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGCTTAACAAGCTTACTCGAGGTCA 367
      381 CGCGTGTACCAAGCTTGATATCCCTAGGCGCGCGCTTAACAAGCTTACTCGAGGTCA 322
Db      368 TTCATATGCTTGAGAAAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 425
      321 TTCATATGCTTGAGAAAGAGTGGGATAGTCCAAATAAACAAGTAAGATTACC 264
```

```

RESULT 5
US-10-213-791-27
; Sequence 27, Application US/10213791
; Publication No. US20030106096A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/10/213,791
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US/09/441,340
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/108,763
; PRIOR FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:expression
; OTHER INFORMATION: cassette comprising a plant promoter linked to an
; OTHER INFORMATION: intron, a sequence encoding an AMPA acetyl
; OTHER INFORMATION: transferase, and a termination sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (28)..(965)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (966)..(1423)
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: (1440)..(1667)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1668)..(2099)
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (2114)..(2369)
; US-10-213-791-27

Query Match      20.3%; Score 101; DB 15; Length 2378;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      325 GATATCCCTAGGCGCGCGCTTAACAAGCTTACTCGAGGTCAATTCATATGCTGAGAG 384
      1 GATATCCCTAGGCGCGCGCTTAACAAGCTTACTCGAGGTCAATTCATATGCTGAGAG 60
Db      385 AGAGTCGGGATAGTCCAAATAAACAAGTAAGATTACC 425
      61 AGAGTCGGGATAGTCCAAATAAACAAGTAAGATTACC 101

RESULT 6
US-10-012-070A-49
; Sequence 49, Application US/10012070A
; Publication No. US20030077801A1
; GENERAL INFORMATION:
; APPLICANT: Hawkes, Timothy
; APPLICANT: Warner, Simon
; APPLICANT: Andrews, Christopher
; APPLICANT: Bachoo, Satvinder
; APPLICANT: Pickerill, Andrew
; TITLE OF INVENTION: Herbicide Resistant Plants
; FILE REFERENCE: 50490/UST
; CURRENT APPLICATION NUMBER: US/10/012,070A
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: PCT/GB00/01573
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 57
```

SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 49  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Oryza sp.  
US-10-012-070A-49

Query Match 14.1%; Score 70; DB 14; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 1.8e-27;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGG 415  
DB 750 TACTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGG 809  
QY 416 TAAGATTACC 425  
DB 810 TAAGATTACC 819

RESULT 7  
US-10-012-013-44  
; Sequence 44; Application US/10012013  
; Publication No. US20030079246A1  
; GENERAL INFORMATION:  
; APPLICANT: Hawkes, Timothy  
; APPLICANT: Warner, Simon  
; APPLICANT: Andrews, Christopher  
; APPLICANT: Bachoo, Savinder  
; APPLICANT: Pickertill, Andrew  
; TITLE OF INVENTION: Herbicide Resistant Plants  
; FILE REFERENCE: 50450/US  
; CURRENT APPLICATION NUMBER: US/10/012,013  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: PCT/GB00/01572  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 44  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Oryza sp.  
US-10-012-013-44

Query Match 14.1%; Score 70; DB 14; Length 1501;  
Best Local Similarity 100.0%; Pred. No. 1.8e-27;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 TACTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGG 415  
DB 750 TACTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGG 809  
QY 416 TAAGATTACC 425  
DB 810 TAAGATTACC 819

RESULT 8  
US-09-991-209-43  
; Sequence 43; Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
; FILE REFERENCE: GC648-2  
; CURRENT APPLICATION NUMBER: US/09/991,209  
; PRIOR FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: US 60/249,608  
; PRIOR FILING DATE: 2000-11-17

NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 1259  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: actin promoter  
US-09-991-209-43

Query Match 13.7%; Score 68; DB 10; Length 1259;  
Best Local Similarity 100.0%; Pred. No. 2.3e-26;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGGTA 417  
DB 16 CTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGGTA 75  
QY 418 AGATTACC 425  
DB 76 AGATTACC 83

RESULT 9  
US-10-678-588A-1/c  
; Sequence 1; Application US/10678588A  
; Publication No. US2005002266A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Jingrui  
; TITLE OF INVENTION: Water-Deficit-TolerantTransgenic Plants  
; FILE REFERENCE: 38-21(52578)C  
; CURRENT APPLICATION NUMBER: US/10/678,588A  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US 60/415,758  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: US 60/425,157  
; PRIOR FILING DATE: 2002-11-08  
; PRIOR APPLICATION NUMBER: US 60/463,787  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 2480  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: transcriptional unit comprising promoter, coding sequence for  
; OTHER INFORMATION: transcription factor of SEQ ID NO:2 and terminator elements  
US-10-678-588A-1

Query Match 13.7%; Score 68; DB 19; Length 2480;  
Best Local Similarity 100.0%; Pred. No. 2.3e-26;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGGTA 417  
DB 2480 CTCGAGTCATTGCTTGAAGAAGAGTCGGATGTCCTCAAAATTAACAAGGTA 2421  
QY 418 AGATTACC 425  
DB 2420 AGATTACC 2413

RESULT 10  
US-09-991-209-36  
; Sequence 36; Application US/09991209  
; Publication No. US20030024009A1  
; GENERAL INFORMATION:  
; APPLICANT: Dunn-Coleman, Nigel  
; APPLICANT: Langdon, Timothy  
; APPLICANT: Morse, Phillip  
; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted

```

; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 5164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pPO10.1 vector
US-09-991-209-36

Query Match      13.7%; Score 68; DB 10; Length 5164;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 417
         |||||
DB      3934 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 3993

QY      418 AGATTACC 425
         |||||
DB      3994 AGATTACC 4001

RESULT 11
US-09-991-209-25
; Sequence 25, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 5277
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTP5-1 vector
US-09-991-209-25

Query Match      13.7%; Score 68; DB 10; Length 5277;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 417
         |||||
DB      4050 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 4109

QY      418 AGATTACC 425
         |||||
DB      4110 AGATTACC 4117

RESULT 12
US-09-991-209-38
; Sequence 38, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:

; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 5295
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pUG4 vector
US-09-991-209-38

Query Match      13.7%; Score 68; DB 10; Length 5295;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 417
         |||||
DB      4068 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 4127

QY      418 AGATTACC 425
         |||||
DB      4128 AGATTACC 4135

RESULT 13
US-09-991-209-27
; Sequence 27, Application US/09991209
; Publication No. US20030024009A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Langdon, Timothy
; TITLE OF INVENTION: Manipulation of the Phenolic Acid
; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted
; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes
; FILE REFERENCE: GC648-2
; CURRENT APPLICATION NUMBER: US/09/991,209
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 60/249,608
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 5327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTP4a2 vector
US-09-991-209-27

Query Match      13.7%; Score 68; DB 10; Length 5327;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      358 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 417
         |||||
DB      3876 CTCGAGTCATTGATGCTTGAGAGAGAGTGGGATGTCCTCAAAATTAACCAAGGTA 3935

QY      418 AGATTACC 425
         |||||
DB      3936 AGATTACC 3943
```

RESULT 14  
 US-09-991-209-19  
 ; Sequence 19, Application US/09991209  
 ; Publication No. US20030024009A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dunn-Coleman, Nigel  
 ; APPLICANT: Langdon, Timothy  
 ; APPLICANT: Morse, Phillip  
 ; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
 ; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
 ; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
 ; FILE REFERENCE: GC648-2  
 ; CURRENT APPLICATION NUMBER: US/09/991,209  
 ; CURRENT FILING DATE: 2002-07-02  
 ; PRIOR APPLICATION NUMBER: US 60/249,608  
 ; PRIOR FILING DATE: 2000-11-17  
 ; NUMBER OF SEQ ID NOS: 97  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 19  
 ; LENGTH: 5337  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: PTU4 vector  
 US-09-991-209-19

Query Match 13.7%; Score 68; DB 10; Length 5337;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-26;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 417  
 |||||  
 DB 4110 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 4169

QY 418 AGATTACC 425  
 |||||  
 DB 4170 AGATTACC 4177

RESULT 15  
 US-09-991-209-23  
 ; Sequence 23, Application US/09991209  
 ; Publication No. US20030024009A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dunn-Coleman, Nigel  
 ; APPLICANT: Langdon, Timothy  
 ; APPLICANT: Morse, Phillip  
 ; TITLE OF INVENTION: Manipulation of the Phenolic Acid  
 ; TITLE OF INVENTION: Content and Digestibility of Plant Cell Walls by Targeted  
 ; TITLE OF INVENTION: Expression of Genes Encoding Cell Wall Degrading Enzymes  
 ; FILE REFERENCE: GC648-2  
 ; CURRENT APPLICATION NUMBER: US/09/991,209  
 ; CURRENT FILING DATE: 2002-07-02  
 ; PRIOR APPLICATION NUMBER: US 60/249,608  
 ; PRIOR FILING DATE: 2000-11-17  
 ; NUMBER OF SEQ ID NOS: 97  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 23  
 ; LENGTH: 5337  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: PTP8-5 vector  
 US-09-991-209-23

Query Match 13.7%; Score 68; DB 10; Length 5337;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-26;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 417  
 |||||  
 DB 4110 CTCGAGTCATTCATATGCTTGAGAGAGAGTCGGGATATGTCCAAAATTAACAAGGTA 4169

QY 418 AGATTACC 425  
 |||||  
 DB 4170 AGATTACC 4177

Search completed: February 9, 2005, 11:12:26  
 Job time : 829.202 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 10, 2005, 07:18:58 ; Search time 5.53136 Seconds  
(without alignments)  
5324.730 Million cell updates/sec

Title: US-10-790-430-11

Perfect score: 18 TGCCTCTCTGCTGACTTC 18

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCUS.COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 18    | 100.0       | 18     | 4     | US-09-872-051-11     |
| 2          | 18    | 100.0       | 18     | 4     | US-09-872-051-8      |
| 3          | 16    | 88.9        | 601    | 4     | US-09-949-016-81715  |
| 4          | 16    | 88.9        | 601    | 4     | US-09-949-016-81717  |
| 5          | 16    | 88.9        | 31391  | 4     | US-09-949-016-14319  |
| 6          | 16    | 88.9        | 125192 | 4     | US-09-949-016-14120  |
| 7          | 15    | 83.3        | 21     | 4     | US-09-949-016-7605   |
| 8          | 15    | 83.3        | 601    | 4     | US-09-949-016-58184  |
| 9          | 15    | 83.3        | 21125  | 4     | US-09-949-016-15108  |
| 10         | 15    | 83.3        | 57761  | 4     | US-09-949-016-13429  |
| 11         | 14    | 77.8        | 20     | 4     | US-09-198-452A-5457  |
| 12         | 14    | 77.8        | 241    | 4     | US-09-016-434-403    |
| 13         | 14    | 77.8        | 356    | 4     | US-09-621-976-9793   |
| 14         | 14    | 77.8        | 459    | 4     | US-09-489-039A-4     |
| 15         | 14    | 77.8        | 601    | 4     | US-09-949-016-19121  |
| 16         | 14    | 77.8        | 601    | 4     | US-09-949-016-27127  |
| 17         | 14    | 77.8        | 601    | 4     | US-09-949-016-35366  |
| 18         | 14    | 77.8        | 601    | 4     | US-09-949-016-72490  |
| 19         | 14    | 77.8        | 601    | 4     | US-09-949-016-102817 |
| 20         | 14    | 77.8        | 601    | 4     | US-09-949-016-102836 |
| 21         | 14    | 77.8        | 601    | 4     | US-09-949-016-160521 |
| 22         | 14    | 77.8        | 601    | 4     | US-09-949-016-203142 |
| 23         | 14    | 77.8        | 601    | 4     | US-09-949-016-203184 |
| 24         | 14    | 77.8        | 1001   | 4     | US-09-671-317-418    |
| 25         | 14    | 77.8        | 1797   | 4     | US-09-891-661-34     |
| 26         | 14    | 77.8        | 2465   | 1     | US-08-421-661-5      |
| 27         | 14    | 77.8        | 2634   | 4     | US-09-949-016-4496   |

|   |    |    |      |       |   |                     |                   |
|---|----|----|------|-------|---|---------------------|-------------------|
| C | 28 | 14 | 77.8 | 2808  | 4 | US-09-917-254-27    | Sequence 27, Appl |
|   | 29 | 14 | 77.8 | 4181  | 1 | US-07-670-611-1     | Sequence 1, Appl1 |
|   | 30 | 14 | 77.8 | 4181  | 1 | US-08-220-674-1     | Sequence 1, Appl1 |
|   | 31 | 14 | 77.8 | 4181  | 1 | US-08-445-186-1     | Sequence 1, Appl1 |
|   | 32 | 14 | 77.8 | 4181  | 1 | US-08-446-549-1     | Sequence 1, Appl1 |
|   | 33 | 14 | 77.8 | 4181  | 2 | US-08-446-550-1     | Sequence 1, Appl1 |
|   | 34 | 14 | 77.8 | 5385  | 4 | US-09-920-804-1     | Sequence 2, Appl1 |
| C | 35 | 14 | 77.8 | 5661  | 3 | US-08-938-105-2     | Sequence 2, Appl1 |
|   | 36 | 14 | 77.8 | 11726 | 4 | US-09-949-016-16968 | Sequence 16968, A |
|   | 37 | 14 | 77.8 | 11726 | 4 | US-09-949-016-16969 | Sequence 16969, A |
| C | 38 | 14 | 77.8 | 12725 | 4 | US-09-949-016-14177 | Sequence 14177, A |
|   | 39 | 14 | 77.8 | 16341 | 4 | US-09-949-016-14579 | Sequence 14579, A |
|   | 40 | 14 | 77.8 | 16341 | 4 | US-09-949-016-14580 | Sequence 14580, A |
| C | 41 | 14 | 77.8 | 20063 | 4 | US-09-949-016-14176 | Sequence 14176, A |
|   | 42 | 14 | 77.8 | 20530 | 4 | US-09-949-016-14394 | Sequence 14394, A |
|   | 43 | 14 | 77.8 | 21784 | 4 | US-09-820-002-3     | Sequence 3, Appl1 |
|   | 44 | 14 | 77.8 | 29927 | 4 | US-09-949-016-11814 | Sequence 11814, A |
|   | 45 | 14 | 77.8 | 29927 | 4 | US-09-949-016-17474 | Sequence 17474, A |

#### ALIGNMENTS

```
RESULT 1
US-09-872-051-11
; Sequence 11, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jimhong
; TITLE OF INVENTION: Corn Event PV-ZMG32 (nK603) and Composition and Methods for Det
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(18)
; OTHER INFORMATION: Zea maize genomic DNA and vector DNA
US-09-872-051-11

Query Match      100.0%; Score 18; DB 4; Length 18;
Best local similarity 100.0%; Pred. No. 0.54;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1  TGCCTCTCTGCTGACTTC 18
Db      1  TGCCTCTCTGCTGACTTC 18

RESULT 2
US-09-872-051-8
; Sequence 8, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
```

```

; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMG32 (nk603) and Composition and Methods for Detecting
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

```

```

Query Match      100.0%; Score 18; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGCTGTTCTGCTGACTT 18
DB      678 TGCTGTTCTGCTGACTT 635

```

```

RESULT 3
US-09-949-016-81715/c
; Sequence 81715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81715
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-81715

```

```

Query Match      88.9%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGCTGTTCTGCTGACT 16
DB      349 TGCTGTTCTGCTGACT 334

```

```

RESULT 4
US-09-949-016-81717/c

```

```

; Sequence 81717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81717
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-81717

```

```

Query Match      88.9%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGCTGTTCTGCTGACT 16
DB      223 TGCTGTTCTGCTGACT 208

```

```

RESULT 5
US-09-949-016-14319
; Sequence 14319, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14319
; LENGTH: 31391
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(31391)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14319

```

```

Query Match      88.9%; Score 16; DB 4; Length 31391;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TGCTGTTCTGCTGACT 16
DB      21183 TGCTGTTCTGCTGACT 21198

```

```

RESULT 6
US-09-949-016-14120/c
; Sequence 14120, Application US/09949016

```

```
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14120
LENGTH: 125192
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(125192)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14120
```

```
Query Match      88.3%; Score 16; DB 4; Length 125192;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGCTGTTCTGCTGACT 16
DB      48907 TGCTGTTCTGCTGACT 48892
```

```
RESULT 7
US-09-422-978-7605
Sequence 7605, Application US/09422978
Patent No. 6537751
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
FILE REFERENCE: GENSET.020CP1
CURRENT APPLICATION NUMBER: US/09/422,978
CURRENT FILING DATE: 1999-10-20
PRIOR APPLICATION NUMBER: 60/298,850
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: US 60/109,732
PRIOR FILING DATE: 1998-11-23
PRIOR APPLICATION NUMBER: US 60/082,614
PRIOR FILING DATE: 1999-04-21
NUMBER OF SEQ ID NOS: 11796
SEQ ID NO 7605
LENGTH: 21
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1..21
OTHER INFORMATION: upstream amplification primer 99-9623 for SEQ 3671,
US-09-422-978-7605
```

```
Query Match      83.3%; Score 15; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CTGTTCTGCTGACTT 17
DB      2 CTGTTCTGCTGACTT 16
```

```
RESULT 8
US-09-949-016-58184
Sequence 58184, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 58184
LENGTH: 601
TYPE: DNA
ORGANISM: Human
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-58184
```

```
Query Match      83.3%; Score 15; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 TGTTCTGCTGACTTT 18
DB      512 TGTTCTGCTGACTTT 526
```

```
RESULT 9
US-09-949-016-15108
Sequence 15108, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15108
LENGTH: 21125
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(21125)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15108
```

```
Query Match      83.3%; Score 15; DB 4; Length 21125;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CTGTTCTGCTGACTT 17
DB      11408 CTGTTCTGCTGACTT 11422
```

RESULT 10

US-09-949-016-13429/c  
; Sequence 13429, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 13429  
; LENGTH: 57761  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(57761)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13429

Query Match 83.3%; Score 15; DB 4; Length 57761;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 TGTCTGCTGACTTT 18  
Db 27356 TGTCTGCTGACTTT 27342

RESULT 11  
US-09-198-452A-5457/c  
; Sequence 5457, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffiths, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 5457  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-5457

Query Match 77.8%; Score 14; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 GTTCTGCTGACTTT 18  
Db 16 GTTCTGCTGACTTT 3

RESULT 12  
US-09-016-434-403/c  
; Sequence 403, Application US/09016434  
; Patent No. 6500338  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 403:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRSTNOT07  
; CLONE: 2121175  
US-09-016-434-403

Query Match 77.8%; Score 14; DB 4; Length 241;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGTCTGCTGCTGA 14  
Db 163 TGTCTGCTGCTGA 150

RESULT 13  
US-09-621-976-9793/c  
; Sequence 9793, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Mline Edwards, J.B.  
; APPLICANT: Jobert, S.  
; TITLE OF INVENTION: Giordano, J.Y.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9793  
; LENGTH: 356  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-9793

Query Match 77.8%; Score 14; DB 4; Length 356;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 4 TGCTGCTGACTT 17  
 |||||  
 DB 85 TGCTGCTGACTT 72

RESULT 14  
 US-09-489-039A-4  
 ; Sequence 4, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; CURRENT FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 4  
 ; LENGTH: 459  
 ; TYPE: DNA  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-4

Query Match 77.8%; Score 14; DB 4; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGCTGCTGA 14  
 |||||  
 DB 258 TGCTGCTGCTGA 271

RESULT 15  
 US-09-949-016-19121/C  
 ; Sequence 19121, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 19121  
 ; LENGTH: 601  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-19121

Query Match 77.8%; Score 14; DB 4; Length 601;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTGCTGCTGA 14  
 |||||  
 DB 553 TGCTGCTGCTGA 540

Search completed: February 10, 2005, 08:50:03  
 Job time : 7.5316 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 10, 2005, 07:18:58 ; Search time 5.5316 Seconds  
(Without alignments)  
5324.730 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 18

Sequence: 1 taccacgcgacacacttc 18

Scoring table: OLIGO NUC  
Gapop 60.0, Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patente NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length  | DB ID                  | Description        |
|------------|-------|-------------|---------|------------------------|--------------------|
| 1          | 18    | 100.0       | 18      | 4 US-09-872-051-10     | Sequence 10, Appl  |
| 2          | 18    | 100.0       | 18      | 4 US-09-872-051-8      | Sequence 8, Appl   |
| 3          | 14    | 77.8        | 174170  | 4 US-09-949-016-14810  | Sequence 14810, A  |
| 4          | 14    | 77.8        | 174170  | 4 US-09-949-016-14811  | Sequence 14811, A  |
| 5          | 14    | 77.8        | 174318  | 4 US-09-949-016-11880  | Sequence 11880, A  |
| 6          | 14    | 77.8        | 174318  | 4 US-09-949-016-14812  | Sequence 14812, A  |
| 7          | 14    | 77.8        | 174318  | 4 US-09-949-016-14813  | Sequence 14813, A  |
| 8          | 13    | 72.2        | 296     | 4 US-09-313-294A-4452  | Sequence 4452, Ap  |
| 9          | 13    | 72.2        | 321     | 4 US-09-270-767-29026  | Sequence 29026, A  |
| 10         | 13    | 72.2        | 1419    | 4 US-09-270-767-13127  | Sequence 13127, A  |
| 11         | 13    | 72.2        | 35081   | 2 US-08-752-760A-1     | Sequence 1, Appl   |
| 12         | 13    | 72.2        | 4403765 | 3 US-09-103-840A-2     | Sequence 2, Appl   |
| 13         | 13    | 72.2        | 4411529 | 3 US-09-103-840A-1     | Sequence 1, Appl   |
| 14         | 12    | 66.7        | 84      | 4 US-09-513-999C-15367 | Sequence 15367, A  |
| 15         | 12    | 66.7        | 262     | 4 US-09-513-999C-14094 | Sequence 14094, A  |
| 16         | 12    | 66.7        | 435     | 4 US-09-270-767-6685   | Sequence 6685, Ap  |
| 17         | 12    | 66.7        | 435     | 4 US-09-270-767-21967  | Sequence 21967, A  |
| 18         | 12    | 66.7        | 477     | 4 US-09-270-767-4902   | Sequence 4902, Ap  |
| 19         | 12    | 66.7        | 477     | 4 US-09-270-767-20184  | Sequence 20184, A  |
| 20         | 12    | 66.7        | 548     | 4 US-09-791-105B-31    | Sequence 31, Appl  |
| 21         | 12    | 66.7        | 558     | 4 US-09-902-540-1969   | Sequence 1969, Ap  |
| 22         | 12    | 66.7        | 601     | 4 US-09-949-016-57353  | Sequence 57353, A  |
| 23         | 12    | 66.7        | 601     | 4 US-09-949-016-205395 | Sequence 205395, A |
| 24         | 12    | 66.7        | 691     | 3 US-08-998-416-742    | Sequence 742, App  |
| 25         | 12    | 66.7        | 712     | 3 US-08-998-416-827    | Sequence 827, App  |
| 26         | 12    | 66.7        | 813     | 3 US-08-818-112-3      | Sequence 3, Appl   |
| 27         | 12    | 66.7        | 813     | 3 US-08-818-111-3      | Sequence 3, Appl   |

|    |    |      |      |                       |                   |
|----|----|------|------|-----------------------|-------------------|
| 28 | 12 | 66.7 | 813  | 3 US-09-056-556-3     | Sequence 3, Appl  |
| 29 | 12 | 66.7 | 813  | 4 US-09-072-596-3     | Sequence 3, Appl  |
| 30 | 12 | 66.7 | 1378 | 4 US-09-072-967-3     | Sequence 3, Appl  |
| 31 | 12 | 66.7 | 1378 | 4 US-09-621-377B-1    | Sequence 1, Appl  |
| 32 | 12 | 66.7 | 1378 | 4 US-10-031-893-1     | Sequence 1, Appl  |
| 33 | 12 | 66.7 | 1391 | 2 US-08-813-940-5     | Sequence 5, Appl  |
| 34 | 12 | 66.7 | 1414 | 4 US-09-270-767-11194 | Sequence 11194, A |
| 35 | 12 | 66.7 | 1497 | 4 US-09-270-767-11259 | Sequence 11259, A |
| 36 | 12 | 66.7 | 1539 | 4 US-09-716-865-13    | Sequence 13, Appl |
| 37 | 12 | 66.7 | 1668 | 2 US-08-901-547A-1    | Sequence 1, Appl  |
| 38 | 12 | 66.7 | 1859 | 3 US-07-861-458C-1    | Sequence 1, Appl  |
| 39 | 12 | 66.7 | 2558 | 3 US-08-999-733-3     | Sequence 3, Appl  |
| 40 | 12 | 66.7 | 3093 | 4 US-09-614-221A-108  | Sequence 108, App |
| 41 | 12 | 66.7 | 3483 | 4 US-09-711-164-230   | Sequence 230, App |
| 42 | 12 | 66.7 | 4175 | 1 US-08-306-691B-49   | Sequence 49, Appl |
| 43 | 12 | 66.7 | 4175 | 3 US-08-202-841A-1    | Sequence 1, Appl  |
| 44 | 12 | 66.7 | 4175 | 5 PCT-US93-06251-84   | Sequence 84, Appl |
| 45 | 12 | 66.7 | 9408 | 4 US-09-418-710-14    | Sequence 14, Appl |

## ALIGNMENTS

```

RESULT 1
US-09-872-051-10
; Sequence 10, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine
; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMG732(hk603) and Composition and Methods for Det
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 2000-10-13
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(18)
; OTHER INFORMATION: zea maize plastid DNA and vector DNA
US-09-872-051-10
Query Match 100.0%; Score 18; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TACCACGCGACACACTTC 18
Db 1 TACCACGCGACACACTTC 18
RESULT 2
US-09-872-051-8
; Sequence 8, Application US/09872051
; Patent No. 6825400
; GENERAL INFORMATION:
; APPLICANT: Monsanto Co
; APPLICANT: Behr, Carl
; APPLICANT: Hironaka, Catherine

```

```

; APPLICANT: Heck, Gregory
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: Corn Event PV-ZMG32 (nk603) and Composition and Methods for Detecting
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-21(52258)B
; CURRENT APPLICATION NUMBER: US/09/872,051
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/213,567
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/241,215
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/240,014
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1183)
; OTHER INFORMATION: 1-164 Agrobacterium tumefaciens nos 3' terminator
; OTHER INFORMATION: 165-381 construct vector DNA
; OTHER INFORMATION: 382-686 Zea maize plastid genes, rps11 and rpoA
; OTHER INFORMATION: 687-1183 Zea maize genomic DNA
US-09-872-051-8

```

```

Query Match          100.0%; Score 18; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TACCGCGACACTT 18
    |||||
Db 373 TACCGCGACACTT 390

```

```

RESULT 3
US-09-949-016-14810
; Sequence 14810, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14810
; LENGTH: 174170
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174170)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14810

```

```

Query Match          77.8%; Score 14; DB 4; Length 174170;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CACCGGACACTT 17
    |||||
Db 86907 CACCGGACACTT 86920

```

```

RESULT 4
US-09-949-016-14811
; Sequence 14811, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14811
; LENGTH: 174170
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174170)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14811

```

```

Query Match          77.8%; Score 14; DB 4; Length 174170;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 CACCGGACACTT 17
    |||||
Db 86907 CACCGGACACTT 86920

```

```

RESULT 5
US-09-949-016-11880
; Sequence 11880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11880
; LENGTH: 174318
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174318)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11880

```

```

Query Match          77.8%; Score 14; DB 4; Length 174318;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 CACGGCAGACACTT 17  
| | | | | | | | | |  
Db 87055 CACGGCAGACACTT 87068

RESULT 6  
US-09-949-016-14812

; Sequence 14812, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14812  
; LENGTH: 174318  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(174318)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14812

Query Match 77.8%; Score 14; DB 4; Length 174318;  
Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGGCAGACACTT 17  
| | | | | | | | | |  
Db 87055 CACGGCAGACACTT 87068

RESULT 7  
US-09-949-016-14813  
; Sequence 14813, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14813  
; LENGTH: 174318  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(174318)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14813

Query Match 77.8%; Score 14; DB 4; Length 174318;

Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGGCAGACACTT 17  
| | | | | | | | | |  
Db 87055 CACGGCAGACACTT 87068

RESULT 8

US-09-313-294A-4452  
; Sequence 4452, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalugudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 4452  
; LENGTH: 296  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700348447H1  
US-09-313-294A-4452

Query Match 72.2%; Score 13; DB 4; Length 296;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACCACGGCAGACACA 14  
| | | | | | | | | |  
Db 260 ACCACGGCAGACACA 272

RESULT 9  
US-09-270-767-29026/c  
; Sequence 29026, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29026  
; LENGTH: 321  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-29026

Query Match 72.2%; Score 13; DB 4; Length 321;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCAACGGCAGACAC 15  
| | | | | | | | | |  
Db 62 CCAACGGCAGACAC 50

RESULT 10  
US-09-270-767-13127/c  
; Sequence 13127, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
;; FILE REFERENCE: File Reference: 7326-094  
;; CURRENT APPLICATION NUMBER: US/09/270,767  
;; CURRENT FILING DATE: 1999-03-17  
;; NUMBER OF SEQ ID NOS: 62517  
;; SOFTWARE: Patentin Ver. 2.0  
;; SEQ ID NO 13127  
;; LENGTH: 1419  
;; TYPE: DNA  
;; ORGANISM: Drosophila melanogaster  
US-09-270-767-13127

Query Match 72.2%; Score 13; DB 4; Length 1419;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCGCGCACACTTC 15  
DB 536 CCGCGCACACTTC 524

RESULT 11  
US-08-752-760A-1  
; Sequence 1, Application US/08752760A  
; Patent No. 5877011  
; GENERAL INFORMATION:  
; APPLICANT: Armentano, Donna  
; APPLICANT: Gregory, Richard J.  
; APPLICANT: Smith, Alan E.  
; TITLE OF INVENTION: CHIMERIC ADENOIRAL VECTORS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Baker & Botts, L.L.P.  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,760A  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seide, Rochelle K  
; REGISTRATION NUMBER: 32,300  
; REFERENCE/DOCKET NUMBER: A31385  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-705-5000  
; TELEFAX: 212-705-5020  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35081 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-752-760A-1

Query Match 72.2%; Score 13; DB 2; Length 35081;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CCGCGCACACTTC 18  
DB 6154 CCGCGCACACTTC 6166

RESULT 12  
US-09-103-840A-2/C  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 72.2%; Score 13; DB 3; Length 4403765;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ACGCGACACTT 17  
DB 1224061 ACGCGACACTT 1224049

RESULT 13  
US-09-103-840A-1/C  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 72.2%; Score 13; DB 3; Length 4411529;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ACGCGACACTT 17  
DB 1224533 ACGCGACACTT 1224521

RESULT 14  
US-09-513-999C-15367/C  
; Sequence 15367, Application US/09513999C  
; Patent No. 6783961

```

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15367
; LENGTH: 84
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-15367

```

```

Query Match      66.7%; Score 12; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      7 GCGACACACTTC 18
        |||||
Db      16 GCGACACACTTC 5

```

```

RESULT 15
US-09-513-999C-14094/C
; Sequence 14094, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14094
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 35
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 39
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 41
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 50
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 262
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-14094

```

```

Query Match      66.7%; Score 12; DB 4; Length 262;
Best Local Similarity 100.0%; Pred. No. 2e+02;

```

```

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2 ACCACGCGACAC 13
        |||||
Db      199 ACCACGCGACAC 188

```

```

Search completed: February 10, 2005, 08:50:01
Job time : 18.5314 secs

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using BW model

Run on: February 10, 2005, 07:47:34 ; Search time 111.027 Seconds  
(without alignment)  
7855.666 Million cell updates/sec

Title: US-10-790-430-10

Perfect score: 18

Sequence: 1 Taccacgcgcacacttc 18

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : GenEmb1.\*  
1: gb\_Da.\*  
2: gb\_hlg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 18    | 100.0       | 18     | AX342371    | Sequence           |
| 2          | 18    | 100.0       | 183    | AX342369    | Sequence           |
| 3          | 17    | 94.4        | 865    | HSB325002   | Homo sapi          |
| 4          | 16    | 88.9        | 98319  | AC090523    | Caenorhabd         |
| 5          | 15    | 83.3        | 363    | BV074353    | Sequence           |
| 6          | 15    | 83.3        | 747    | AX065627    | Sequence           |
| 7          | 15    | 83.3        | 69794  | AC013773    | Homo sapi          |
| 8          | 15    | 83.3        | 89474  | AC013802    | Homo sapi          |
| 9          | 15    | 83.3        | 110000 | AP006502_01 | Continuation (2 of |
| 10         | 15    | 83.3        | 170777 | AC148444    | Rhinoloph          |
| 11         | 15    | 83.3        | 299587 | AC091307    | Mus muscu          |
| 12         | 15    | 83.3        | 309400 | AX127153    | Sequence           |
| 13         | 15    | 83.3        | 325651 | AP005283    | Corynebact         |
| 14         | 15    | 83.3        | 349115 | BX927156    | Corynebact         |
| 15         | 15    | 83.3        | 349980 | AX127152    | Sequence           |
| 16         | 14    | 77.8        | 435    | CO499894    | Sequence           |
| 17         | 14    | 77.8        | 435    | CO508863    | Sequence           |
| 18         | 14    | 77.8        | 435    | AX308198    | Sequence           |
| 19         | 14    | 77.8        | 437    | CO469529    | Sequence           |

|   |    |    |      |        |    |             |                    |
|---|----|----|------|--------|----|-------------|--------------------|
| c | 20 | 14 | 77.8 | 484    | 6  | CO478697    | Sequence           |
| c | 21 | 14 | 77.8 | 1889   | 9  | AF383169    | Pan trogl          |
| c | 22 | 14 | 77.8 | 2000   | 6  | AX655473    | Sequence           |
| c | 23 | 14 | 77.8 | 3164   | 5  | BC044418    | Danio rer          |
| c | 24 | 14 | 77.8 | 5813   | 6  | CO580220    | Sequence           |
| c | 25 | 14 | 77.8 | 1111   | 1  | AE007286    | Simorhizo          |
| c | 26 | 14 | 77.8 | 29989  | 3  | U40059      | Caenorhabd         |
| c | 27 | 14 | 77.8 | 71018  | 2  | AC131498    | Lytechinu          |
| c | 28 | 14 | 77.8 | 90027  | 2  | AC079522    | Mus muscu          |
| c | 29 | 14 | 77.8 | 96438  | 2  | AC150979    | Medicago           |
| c | 30 | 14 | 77.8 | 97538  | 2  | AC017943    | Drosophi1          |
| c | 31 | 14 | 77.8 | 105805 | 2  | AC022280    | Drosophi1          |
| c | 32 | 14 | 77.8 | 110000 | 1  | AE017283_09 | Continuation (10 o |
| c | 33 | 14 | 77.8 | 110000 | 2  | LMFCHR36_10 | Continuation (11 o |
| c | 34 | 14 | 77.8 | 112673 | 2  | AY648026    | Ovis arie          |
| c | 35 | 14 | 77.8 | 118020 | 2  | AC134824    | Medicago           |
| c | 36 | 14 | 77.8 | 129402 | 2  | AC142062    | Rattus sat         |
| c | 37 | 14 | 77.8 | 131054 | 8  | AP003341    | Oryza sat          |
| c | 38 | 14 | 77.8 | 135611 | 14 | AF169823    | Spodopter          |
| c | 39 | 14 | 77.8 | 144023 | 9  | AC083928    | Homo sapi          |
| c | 40 | 14 | 77.8 | 156889 | 2  | AC004586    | Homo sapi          |
| c | 41 | 14 | 77.8 | 158084 | 8  | AP003492    | Oryza sat          |
| c | 42 | 14 | 77.8 | 158594 | 2  | AC087768    | Homo sapi          |
| c | 43 | 14 | 77.8 | 165083 | 10 | AC120554    | Homo sapi          |
| c | 44 | 14 | 77.8 | 165497 | 5  | BX511107    | Zebrafish          |
| c | 45 | 14 | 77.8 | 170939 | 3  | AC007757    | Drosophi1          |

## ALIGNMENTS

|                       |  |                                    |               |            |                 |
|-----------------------|--|------------------------------------|---------------|------------|-----------------|
| RESULT 1              | AX342371   | 18 bp                              | DNA           | linear     | PAT 12-JAN-2002 |
| LOCUS                 | AX342371   | Sequence 10 from Patent EP1167531. |               |            |                 |
| DEFINITION            | AX342371   | AX342371.1 GI:18151814             |               |            |                 |
| ACCESSION             | AX342371   |                                    |               |            |                 |
| VERSION               | AX342371.1   | GI:18151814                        |               |            |                 |
| KEYWORDS              |  |                                    |               |            |                 |
| SOURCE                |  |                                    |               |            |                 |
| ORGANISM              |  |                                    |               |            |                 |
| REFERENCE             | 1  |                                    |               |            |                 |
| AUTHORS               | Behr, C.F., Hironaka, C., Heck, G.R. and You, J.                                       |                                    |               |            |                 |
| TITLE                 | Corn transformant pV-zmgf32 (nke03) and compositions and methods for detection thereof |                                    |               |            |                 |
| JOURNAL               | Patent: EP 1167531-A 10 02-JAN-2002;   |                                    |               |            |                 |
| FEATURES              | Monsanto Technology LLC (US)   |                                    |               |            |                 |
| SOURCE                | Location/Qualifiers  |                                    |               |            |                 |
|                       | 1..18  |                                    |               |            |                 |
|                       | /organism="synthetic construct"  |                                    |               |            |                 |
|                       | /mol_type="unassigned DNA"   |                                    |               |            |                 |
|                       | /db_xref="taxon:32630"   |                                    |               |            |                 |
|                       | /note="zea maize plastid DNA and vector DNA"   |                                    |               |            |                 |
| ORIGIN                |  |                                    |               |            |                 |
| Query Match           | 100.0%;  | Score 18;                          | DB 6;         | Length 18; |                 |
| Best Local Similarity | 100.0%;  | Pred. No. 0.57;                    |               |            |                 |
| Matches               | 18;  | Conservative 0;                    | Mismatches 0; | Indels 0;  | Gaps 0;         |
| Db                    | 1 TACCACGCGACACTTC 18  |                                    |               |            |                 |
|                       | 1 TACCACGCGACACTTC 18  |                                    |               |            |                 |
| RESULT 2              | AX342369   | 1183 bp                            | DNA           | linear     | PAT 12-JAN-2002 |
| LOCUS                 | AX342369   | Sequence 8 from Patent EP1167531.  |               |            |                 |
| DEFINITION            | AX342369   | AX342369.1 GI:18151812             |               |            |                 |
| ACCESSION             | AX342369   |                                    |               |            |                 |
| VERSION               | AX342369.1   | GI:18151812                        |               |            |                 |
| KEYWORDS              |  |                                    |               |            |                 |
| SOURCE                |  |                                    |               |            |                 |
| ORGANISM              |  |                                    |               |            |                 |
|                       | synthetic construct  |                                    |               |            |                 |
|                       | synthetic construct  |                                    |               |            |                 |

other sequences; artificial sequences.

REFERENCE 1  
Behr, C.F., Hironaka, C., Heck, G.R. and You, J.  
TITLE Corn transformat pv-zmgc32 (mk603) and compositions and methods for detection thereof  
JOURNAL Patent: EP 1167531-A 8 02-JAN-2002;  
Monsanto Technology LLC (US)  
FEATURES location/Qualifiers  
source 1. 1183  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="1-164 Agrobacterium tumefaciens nos 3' terminator 165-381 construct vector DNA 382-686 Zea mays plasmid genes, rps11 and rps4 687-1183 Zea mays genomic DNA"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACCACGGACACACTTC 18  
|||||  
373 TACCACGGACACACTTC 390

RESULT 3  
HSA325002 865 bp DNA linear PRI 18-JUL-2002  
LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone  
DEFINITION NL1-DK11R.  
ACCESSION AJ325002  
VERSION AJ325002.1 GI:15869396  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 865)  
Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasna, S.M., Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V., Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I., Kisselev, L.L., Wasserman, W., Wahlestedt, C. and Zabarovsky, E.R.  
TITLE Not1 flanking sequences: a tool for gene discovery and verification of the human genome  
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)  
MEDLINE 22131767  
PUBMED 12136098  
REFERENCE 2 (bases 1 to 865)  
Zabarovsky, E.R.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden

FEATURES  
source location/Qualifiers  
1. 865  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3606"  
/clone="NL1-DK11R"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 865;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACCACGGACACACTTC 18  
|||||  
546 ACCACGGACACACTTC 562

RESULT 4  
AC090523

LOCUS AC090523 98319 bp DNA linear INV 27-FEB-2001  
DEFINITION Caenorhabditis briggsae cosmid CB015N01, complete sequence.  
ACCESSION AC090523  
VERSION AC090523.1 GI:13129526  
KEYWORDS HTG.  
SOURCE Caenorhabditis briggsae  
ORGANISM Caenorhabditis briggsae  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 98319)  
The C. briggsae Genome Sequencing Project  
TITLE Washington University Genome Sequencing Center.  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 98319)  
Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (27-FEB-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63110, USA  
e-mail: jpeithewatson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES  
source location/Qualifiers  
1. 98319  
/organism="Caenorhabditis briggsae"  
/mol\_type="genomic DNA"  
/strain="Gujarat G16"  
/db\_xref="taxon:6238"  
/clone="CB015N01"  
/product="cDNA-Len"  
/note="codon recognized: CUG"  
/product="cRNA-Pro"  
/note="codon recognized: CCA"

ORIGIN

Query Match 88.9%; Score 16; DB 3; Length 98319;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGGACACACTTC 18  
|||||  
20762 CCACGGACACACTTC 20777

RESULT 5  
BV074353 363 bp DNA linear STS 31-MAY-2003  
LOCUS S212P609FBI.T0 CZECHII/BI Mus musculus STS genomic, sequence tagged site.  
DEFINITION BV074353  
ACCESSION BV074353  
VERSION BV074353.1 GI:31190148  
KEYWORDS STS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 363)  
Wade, C.M., Kulbokas, E.J., III, Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K. and Daly, M.J.  
TITLE The mosaic structure of variation in the laboratory mouse genome  
JOURNAL Nature 420 (6915), 574-578 (2002)  
MEDLINE 22354684  
PUBMED 12466852  
COMMENT

Contact: Kerstin Lindblad-Toh  
 Whitehead Institute for Biomedical Research, Center for Genome  
 Research  
 320 Charles Street, Cambridge, MA 02141, USA  
 Tel: 6172521477  
 Fax: 6172580903  
 Email: kersti@genome.wi.mit.edu  
 Primer A: None  
 Primer B: None  
 STS size: 363

## Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
 were generated from 129S1/SvImt, C3H/HeJ, and BALB/cByJ. The WGS  
 reads were placed uniquely on the MGSv3 C57BL/6J assembly and SNP  
 detection was carried out by SSAPHA-SNP. 225,000 reads were  
 annotated  
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J  
 and the strain from which the particular read came. The validation  
 rate for these SNPs was estimated at approximately 98%.

## FEATURES

## source

Location/Qualifiers

1..363  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /map="16 27-306 28672947-28672668"  
 /clone\_lib="CZEBCH1/E1"  
 <1..>363

## ORIGIN

## STS

Query Match 83.3%; Score 15; DB 11; Length 363;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

4 CACGGCAGACACTTC 18  
 |||||  
 190 CACGGCAGACACTTC 204

## RESULT 6

AX065627/c 747 bp DNA linear PAT 24-JAN-2001  
 LOCUS AX065627  
 DEFINITION Sequence 753 from Patent WO0100844.  
 ACCESSION AX065627  
 VERSION AX065627.1 GI:12543339

## SOURCE

## ORGANISM

Corynebacterium glutamicum  
 Corynebacterium glutamicum  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

## REFERENCE

1 Pompejus, M., Kroege, B., Schroeder, H., Zelder, O. and Habermann, G.  
 corynebacterium glutamicum genes encoding proteins involved in  
 carbon metabolism and energy production  
 Patent: WO 0100844-A 753 04-JAN-2001;

## JOURNAL

BASF AKTENGESCHLAF (DE)

## FEATURES

## source

1..747  
 /organism="Corynebacterium glutamicum"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1718"  
 101..727  
 /note="unamed protein product; RXN02554"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAC2603.1"  
 /db\_xref="GI:12543340"

## CDS

/translation="MSHTKPSIALIGAGRVSSLSARVAAGYEVKAGSGAVDIAL  
 TAILMPCAVSTADQAKADIVFLAHLKFRSVNPATLGRKIVIDTMHWYPVNG  
 ELERIDDPSTSEILAEFFAGSTWVSFNIIGYHEIDAGTGRAIAVATDVADAGA  
 OVAOLISFGVPLNIGALENGRILEPQEAFGAHLNDSLELVNR"

## ORIGIN

Query Match 83.3%; Score 15; DB 6; Length 747;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

4 CACGGCAGACACTTC 18  
 |||||  
 21 CACGGCAGACACTTC 7

## Db

## RESULT 7

AC013773/c 69794 bp DNA linear HTG 13-JUL-2000  
 LOCUS AC013773  
 DEFINITION Homo sapiens clone RP11-10A24, LOW-PASS SEQUENCE SAMPLING.  
 ACCESSION AC013773  
 VERSION AC013773.2 GI:9123806  
 KEYWORDS HTG; HTGS; PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

## REFERENCE

1 (bases 1 to 69794)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.

## AUTHORS

Unpublished

## JOURNAL

2 (bases 1 to 69794)

## REFERENCES

Baird, J., Berra, N., Beckerly, R., Boguslavsky, L., Bouhgalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., Dearlano, K., Dewar, K., Domini, M., Donelan, L., Doyle, M.,  
 Ferreira, P., Fitzhugh, M., Forrest, C., Funke, R., Gage, D.,  
 Galagan, J., Gardyna, S., Grant, G., Hages, B., Heatford, A., Horton, L.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Lehoczy, J., Liu, C., Locke, K., MacDonald, P., Marquis, N.,  
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Jul 13, 2000 this sequence version replaced gi:642574.

## All repeats were identified using RepeatMasker:

Smil, A. F. A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

## Center code: WtBR

Web site: http://www-seq.wi.mit.edu

## Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

## Center Project name: L3033

Center clone name: 10\_A\_24

NOTE: This record contains 77 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1 792: contig of 792 bp in length  
 \* 793 892: gap of 100 bp  
 \* 893 1698: contig of 806 bp in length  
 \* 1699 1798: gap of 100 bp  
 \* 1799 2602: contig of 804 bp in length  
 \* 2603 2702: gap of 100 bp  
 \* 2703 3518: contig of 816 bp in length

3519 3618: gap of 100 bp  
3619 4415: contig of 797 bp in length  
4416 4515: gap of 100 bp  
4516 5385: contig of 870 bp in length  
5386 5485: gap of 100 bp  
5486 6305: contig of 820 bp in length  
6306 6405: gap of 100 bp  
6406 7233: contig of 828 bp in length  
7234 7333: gap of 100 bp  
7334 8139: contig of 806 bp in length  
8140 8239: gap of 100 bp  
8240 9026: contig of 787 bp in length  
9027 9126: gap of 100 bp  
9127 9933: contig of 806 bp in length  
9933 10033: gap of 100 bp  
10033 10840: contig of 808 bp in length  
10841 10940: gap of 100 bp  
10941 11753: contig of 813 bp in length  
11754 11853: gap of 100 bp  
11854 12657: contig of 804 bp in length  
12658 12757: gap of 100 bp  
12758 13574: contig of 817 bp in length  
13575 13674: gap of 100 bp  
13675 14456: contig of 782 bp in length  
14457 14556: gap of 100 bp  
14557 15353: contig of 797 bp in length  
15354 15453: gap of 100 bp  
15454 16298: contig of 845 bp in length  
16299 16398: gap of 100 bp  
16399 17218: contig of 820 bp in length  
17219 17318: gap of 100 bp  
17319 18116: contig of 798 bp in length  
18117 18215: gap of 100 bp  
18216 19012: contig of 796 bp in length  
19013 19112: gap of 100 bp  
19113 19912: contig of 800 bp in length  
19913 20012: gap of 100 bp  
20012 20835: contig of 823 bp in length  
20836 20935: gap of 100 bp  
20936 21731: contig of 796 bp in length  
21732 21831: gap of 100 bp  
21832 22603: contig of 772 bp in length  
22604 22703: gap of 100 bp  
22704 23507: contig of 804 bp in length  
23508 23607: gap of 100 bp  
23608 24420: contig of 813 bp in length  
24421 24520: gap of 100 bp  
24521 25341: contig of 821 bp in length  
25342 25441: gap of 100 bp  
25442 26235: contig of 794 bp in length  
26236 26335: gap of 100 bp  
26336 27114: contig of 779 bp in length  
27115 27214: gap of 100 bp  
27216 27997: contig of 783 bp in length  
27998 28097: gap of 100 bp  
28098 28923: contig of 826 bp in length  
28924 29023: gap of 100 bp  
29024 29844: contig of 821 bp in length  
29845 29944: gap of 100 bp  
29945 30746: contig of 802 bp in length  
30747 30846: gap of 100 bp  
30847 31663: contig of 817 bp in length  
31664 31763: gap of 100 bp  
31764 32575: contig of 812 bp in length  
32576 33491: contig of 816 bp in length  
33492 33591: gap of 100 bp  
33592 34407: contig of 816 bp in length  
34408 34507: gap of 100 bp  
34509 35309: contig of 802 bp in length  
35310 35409: gap of 100 bp  
35410 36196: contig of 787 bp in length  
36197 36296: gap of 100 bp

36297 37109: contig of 813 bp in length  
37110 37209: gap of 100 bp  
37210 38023: contig of 813 bp in length  
38023 38122: gap of 100 bp  
38123 38917: contig of 795 bp in length  
38918 39017: gap of 100 bp  
39018 39838: contig of 821 bp in length  
39839 39938: gap of 100 bp  
39939 40723: contig of 784 bp in length  
40724 40822: gap of 100 bp  
40823 41610: contig of 788 bp in length  
41611 41710: gap of 100 bp  
41711 42548: contig of 838 bp in length  
42549 42648: gap of 100 bp  
42649 43469: contig of 821 bp in length  
43470 43569: gap of 100 bp  
43570 44387: contig of 818 bp in length  
44388 44487: gap of 100 bp  
44488 45308: contig of 821 bp in length  
45309 45408: gap of 100 bp  
45409 46219: contig of 811 bp in length  
46220 46320: gap of 100 bp  
46321 47123: contig of 804 bp in length  
47124 47223: gap of 100 bp  
47224 48019: contig of 796 bp in length  
48020 48119: gap of 100 bp  
48120 48955: contig of 837 bp in length  
48956 49056: gap of 100 bp  
49057 49826: contig of 770 bp in length  
49827 49926: gap of 100 bp  
49927 50740: contig of 814 bp in length  
50741 50840: gap of 100 bp  
50841 51638: contig of 798 bp in length  
51639 51738: gap of 100 bp  
51739 52551: contig of 813 bp in length  
52552 52652: gap of 100 bp  
52653 53463: contig of 812 bp in length  
53464 53563: gap of 100 bp  
53564 54355: contig of 793 bp in length  
54356 54456: gap of 100 bp  
54457 55285: contig of 829 bp in length  
55286 55385: gap of 100 bp  
55386 56164: contig of 779 bp in length  
56165 56264: gap of 100 bp  
56265 57080: contig of 816 bp in length  
57081 57180: gap of 100 bp  
57181 57964: contig of 784 bp in length  
57965 58064: gap of 100 bp  
58065 58916: contig of 852 bp in length  
58917 59016: gap of 100 bp  
59017 59868: contig of 852 bp in length  
59869 59968: gap of 100 bp  
59969 60728: contig of 760 bp in length  
60729 60828: gap of 100 bp  
60829 61631: contig of 802 bp in length  
61632 61730: gap of 100 bp  
61731 62566: contig of 836 bp in length  
62567 62666: gap of 100 bp  
62667 63483: contig of 817 bp in length  
63484 63583: gap of 100 bp  
63584 64376: contig of 793 bp in length  
64377 64476: gap of 100 bp  
64477 65284: contig of 808 bp in length  
65285 65384: gap of 100 bp  
65385 66185: contig of 801 bp in length

## Query Match

83.3%; Score 15; DB 2; Length 69794;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCGACACTT 17  
DB 24772 CCACGCGACACTT 24758

RESULT 8  
AC013802  
LOCUS  
DEFINITION  
AC013802 Homo sapiens clone RP11-20M13, LOW-PASS SEQUENCE SAMPLING.  
AC013802.3 GI:9123968  
VERSION  
AC013802.3 HTG: HTGS PHASE0.  
KEYWORDS  
HTG: HTGS PHASE0.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Eureleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 89474)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE  
Homo sapiens, clone RP11-20M13  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 89474)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
Baldwin, J., Barina, N., Becker, R., Boguslavsky, L., Bouckgeater, B.,  
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
Galligan, J., Gaidyna, S., Grant, G., Hagos, B., Harford, A., Horton, L.,  
Howland, J., J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,  
Lehoczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N.,  
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, P.,  
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Sudramanian, A., Talamas, J.,  
Tessier, S., Tittel, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, M. J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 13, 2000 this sequence version replaced gi:6957824.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3932  
Center clone name: 20\_M\_13  
-----  
\* NOTE: This record contains 103 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
1  
\* 726: contig of 726 bp in length  
\* 727 826: gap of 100 bp  
\* 728 827: gap of 100 bp  
\* 1515: contig of 689 bp in length  
\* 1516 1615: gap of 100 bp  
\* 1616 2345: contig of 730 bp in length  
\* 2346 2445: gap of 100 bp  
\* 2446 3154: contig of 705 bp in length  
\* 3155 3254: gap of 100 bp  
\* 3255 3983: contig of 729 bp in length  
\* 3984 4083: gap of 100 bp  
\* 4084 4810: contig of 727 bp in length  
\* 4811 4911: gap of 100 bp  
\* 4911 5625: contig of 715 bp in length  
\* 5625 5725: gap of 100 bp

5726 6448: contig of 723 bp in length  
5649 6449: gap of 100 bp  
7262: contig of 714 bp in length  
7263 7362: gap of 100 bp  
7363 8036: contig of 674 bp in length  
8037 8136: gap of 100 bp  
8137 8854: contig of 718 bp in length  
8855 8954: gap of 100 bp  
8955 9634: contig of 740 bp in length  
9635 9794: gap of 100 bp  
9795 10507: contig of 713 bp in length  
10508 10607: gap of 100 bp  
10608 11333: contig of 726 bp in length  
11334 11433: gap of 100 bp  
11434 12142: contig of 705 bp in length  
12143 12242: gap of 100 bp  
12243 12998: contig of 756 bp in length  
12999 13098: gap of 100 bp  
13099 13826: contig of 728 bp in length  
13827 13926: gap of 100 bp  
13927 14647: contig of 721 bp in length  
14648 14747: gap of 100 bp  
14748 15449: contig of 702 bp in length  
15450 15549: gap of 100 bp  
15550 16249: contig of 700 bp in length  
16250 16349: gap of 100 bp  
16350 17058: contig of 705 bp in length  
17059 17158: gap of 100 bp  
17159 17864: contig of 706 bp in length  
17865 17964: gap of 100 bp  
17965 18697: contig of 733 bp in length  
18698 18797: gap of 100 bp  
18798 19512: contig of 715 bp in length  
19513 19612: gap of 100 bp  
19613 20338: contig of 726 bp in length  
20339 20438: gap of 100 bp  
20439 21177: contig of 739 bp in length  
21178 21277: gap of 100 bp  
21278 22000: contig of 723 bp in length  
22001 22100: gap of 100 bp  
22101 22816: contig of 716 bp in length  
22817 22916: gap of 100 bp  
22917 23627: contig of 711 bp in length  
23628 23727: gap of 100 bp  
23728 24461: contig of 734 bp in length  
24462 24561: gap of 100 bp  
24562 25276: contig of 715 bp in length  
25277 25376: gap of 100 bp  
25377 26097: contig of 721 bp in length  
26098 26197: gap of 100 bp  
26198 26908: contig of 711 bp in length  
26909 27008: gap of 100 bp  
27009 27738: contig of 730 bp in length  
27739 27838: gap of 100 bp  
27839 28561: contig of 723 bp in length  
28562 28661: gap of 100 bp  
28662 29375: contig of 714 bp in length  
29376 29475: gap of 100 bp  
29476 30189: contig of 714 bp in length  
30189 30289: gap of 100 bp  
30290 30997: contig of 708 bp in length  
30998 31097: gap of 100 bp  
31098 31820: contig of 723 bp in length  
31821 31920: gap of 100 bp  
31921 32641: contig of 721 bp in length  
32642 32741: gap of 100 bp  
32742 33469: contig of 728 bp in length  
33469 33569: gap of 100 bp  
33570 34286: contig of 717 bp in length  
34287 34386: gap of 100 bp  
34387 35091: contig of 705 bp in length  
35092 35191: gap of 100 bp  
35192 35901: contig of 710 bp in length



Center clone name: 327G16

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

#### ----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 169390 bases at least Q40  
Consensus quality: 169842 bases at least Q30  
Consensus quality: 170073 bases at least Q20  
Insert size: 13300; agarose-fp  
Insert size: 170177; sum-of-contigs  
Quality coverage: 11.58x in Q20 bases; agarose-fp  
Quality coverage: 9.05x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\*  
\* 1 17243: contig of 17243 bp in length  
\* 17244 17343: gap of unknown length  
\* 17344 24932: contig of 7589 bp in length  
\* 24932 25033: gap of unknown length  
\* 25033 75704: contig of 50672 bp in length  
\* 75704 75804: gap of unknown length  
\* 75804 95337: contig of 19533 bp in length  
\* 95337 95438: gap of unknown length  
\* 95438 101416: contig of 5779 bp in length  
\* 101416 101517: gap of unknown length  
\* 101517 163768: contig of 62252 bp in length  
\* 163768 163869: gap of unknown length  
\* 163869 170777: contig of 6909 bp in length.  
Location/Qualifiers  
1. 170777

#### FEATURES

##### source

/organism="Rhizolophus ferrumequinum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:59479"  
/clone="VMRC7-327G16"  
/clone\_1fb="VMRC7"  
/note="BAC resource: http://bacpac.chori.org/"  
1. 59786  
/note="clone overlaps with GenBank Accession Number  
AC148387 clone VMRC7-122D13 (center project name gc1)"  
misc\_feature  
1. 17243  
/note="assembly\_fragment  
clone\_end:other  
vector\_side:left"  
17344..24932  
misc\_feature  
/note="assembly\_fragment"  
25033..75704  
misc\_feature  
/note="assembly\_fragment"  
75805..95337  
misc\_feature  
/note="assembly\_fragment"  
95438..101416  
misc\_feature  
/note="assembly\_fragment"  
101517..163768  
misc\_feature  
/note="assembly\_fragment"

misc\_feature 11659..170777  
/note="clone overlaps with GenBank Accession Number  
AC148388 clone VMRC7-144D23 (center project name gc1)"  
misc\_feature 163869..170777  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:right"

#### ORIGIN

Query Match 83.3%; Score 15; DB 2; Length 170777;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 3 CCACGGCAGCAGCTT 17  
|||||  
DB 79415 CCACGGCAGCAGCTT 79429

#### RESULT 11

AC091307/c

LOCUS

DEFINITION

AC091307

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

O'Neill, D., Oliver, J., Peterson, K., Phunhahang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spenser, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. Direct Submission

Submitted (06-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA  
On Dec 11, 2003 this sequence version replaced gi:24182398.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WITBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L13257

Center clone name: 79\_E\_7

NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 74083: contig of 74083 bp in length  
\* 74084 74183: gap of 100 bp  
\* 74184 110439: contig of 36256 bp in length  
\* 110440 110539: gap of 100 bp  
\* 110540 135499: contig of 24960 bp in length  
\* 135500 135599: gap of 100 bp  
\* 135600 144168: contig of 8569 bp in length  
\* 144169 144268: gap of 100 bp  
\* 144269 177573: contig of 33305 bp in length  
\* 177574 177673: gap of 100 bp  
\* 177674 184331: contig of 6558 bp in length  
\* 184332 205236: contig of 20905 bp in length  
\* 205237 261172: gap of 100 bp  
\* 261173 261272: gap of 100 bp  
\* 261273 280973: contig of 19701 bp in length  
\* 280974 299587: gap of 100 bp  
\* 299588 299587: contig of 18514 bp in length.

FEATURES  
source location/Qualifiers  
1..299587

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="16"

/map="16"

/clone="RP23-79E7"

/clone\_lib="RPCT-23 Female Mouse BAC"

ORIGIN

Query Match 83.3%; Score 15; DB 2; Length 299587;  
Best Local Similarity 100.0%; Pred. No. 22;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
|||||

Db 162261 CACGCGACACACTTC 162247

RESULT 12  
AX127153/c AX127153 309400 bp DNA linear PAT 11-MAY-2001

DEFINITION Sequence 7069 from Patent EP1108790.  
ACCESSION AX127153  
VERSION AX127153.1 GI:14041141

KEYWORDS  
SOURCE  
ORGANISM  
Corynebacterium glutamicum  
Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacterium.

REFERENCE

AUTHORS

TITLE

JOURNAL

KEYWORDS

SOURCE

FEATURES

source

1..309400  
/organism="Corynebacterium glutamicum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1718"  
/note="Seq 1 to 1000 (3,309,400) split in 11, seq 7069  
3,000,001 3,309,400 309,400"

ORIGIN

Query Match 83.3%; Score 15; DB 6; Length 309400;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
|||||

Db 28188 CACGCGACACACTTC 28174

RESULT 13

AP005283/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL





/product="Zn-dependent alcohol dehydrogenases"  
/protein\_id="BAC00201.1"  
/db\_xref="GI:21325580"  
/translation="MTTAAAPREFNAVVEKRGHVTVKDDLPKPGPHQALVKTUTSG  
ICHTDHALGEGDVPKPEPPVPGHEGVEVLEGSDHDKVDIVGNALMSAGT  
CEVCTIGRETFQCEAEYGYTQNGSPQYMLVDRYARLPDGVYLEAAFLICAVT  
VYVALKVSERPGQFMVYISGVGLGHLAVOYAAAMGRVIAVDIADKLRLARHGAE  
FTVARNEDSGEAVQKTYNIGGAGHVLVTAVHEAALFGALDMARAGTIVFGLPPGEF  
PAVFNIVFKGLTIRGSLVGRQDLAEALDPFAGLLKPTVSECSLDEVNGLDRMRN  
GKIDGRAVIRP"  
13402..13620  
/gene="Cg12808"  
13402..13620  
/gene="Cg12808"  
/codon\_start=1

Query Match 83.3%; Score 15; DB 1; Length 325651;  
Best Local. Similarly 100.0%; Pred. No. 22;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACGCGACACACTTC 18  
Db 44439 CACGCGACACACTTC 44425

RESULT 14  
BX927156/c  
LOCUS BX927156 349115 bp DNA linear BCT 10-JUN-2004  
DEFINITION Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5,  
complete genome; segment 9/10.  
ACCESSION BX927156 BX927147  
VERSION BX927156.1 GI:41326831  
KEYWORDS complete genome.  
SOURCE Corynebacterium glutamicum ATCC 13032  
ORGANISM Corynebacterium glutamicum ATCC 13032  
REFERENCE Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
1 (bases 1 to 349115)  
AUTHORS Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M.,  
Burkovski, A., Duesch, N., Eggeling, L., Ekmann, B. J., Gaigalat, L.,  
Goessmann, A., Hartmann, M., Huttmacher, K., Kramer, R., Linke, B.,  
McHardy, A. C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A.,  
Rey, D. A., Ruckert, C., Rupp, O., Sahm, H., Wendisch, V. F., Wiegand, I.  
and Tauch, A.  
TITLE The complete Corynebacterium glutamicum ATCC 13032 genome sequence  
and its impact on the production of L-aspartate-derived amino acids  
and vitamins  
JOURNAL J. Biotechnol. 104 (1-3), 5-25 (2003)  
MEDLINE 22830012  
PUBMED 12948626  
REFERENCE 2 (bases 1 to 349115)  
AUTHORS Kalinowski, J.  
TITLE Direct Submission  
SUBMITTED (21-JUN-2004) Joern Kalinowski, Institut fuer  
Genomforschung, Universitaet Bielefeld; Universitaetsstrasse 25,  
3615 Bielefeld, Germany  
E-mail: Joern.Kalinowski@cebtec.uni-bielefeld.de  
COMMENT This sequence was accomplished by collaboration between Degussa AG  
and Bielefeld University.  
J. Biol. Chem. 276 (18):12948-12954, 2001.  
BX927151.1:51..349459, BX927152.1:51..349799, BX927153.1:51..349584,  
BX927154.1:51..349575, BX927155.1:51..349136, BX927156.1:51..349115,  
BX927157.1:51..140057.  
FEATURES  
Location/Qualifiers  
1..349115  
/organism="Corynebacterium glutamicum ATCC 13032"  
/mol\_type="genomic DNA"  
/strain="DSM 20300 = ATCC 13032"  
/db\_xref="taxon:196627"  
/note="IS fingerprint type: 4-5"  
307..1917  
/locus\_tag="cg2937"

307..1917  
/locus\_tag="cg2937"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type dipeptide/oligopeptide/nickel transport  
system, secreted component"  
/protein\_id="CAF20677.1"  
/db\_xref="GI:41326832"  
/translation="MSTIYIRRNIFLRATGILGVAAGIGATLAACAPDNTSGSTSTA  
AGGTANESGTTTAAISYELGTGVDPMTTSTALTVAAAMHTLEGTEIDPATGVYA  
ALASALPSADATSLDKLADGAPFNHGDVADVVEADSPFANPAGSGPYKCTD  
KSVYTKODTIVTIDLDVATGIISERLAVYIKYKPYVEADAGFPANPAGSGPYKCTD  
NGASKVYKPERNDYNGPPPARAKMEWQIIPDASTRINSLQSGSTMALDVPYLSIP  
QLRATSTVESVQCGFLFAMFNCSEGNPPENRQAPFLYALDKIKYKTMGSDYATP  
ATSFVQKEHPNYQASTVYSLDADKAKALFAETGILSLILCTDHDVYNCPILQES  
LALGIVNSFTERRKSADVYNTTIGRKEADVDVIAPODPEVFGNDPLRMRYAGDVW  
TDSRMKKSSESVDYQVNIIEGRKEATDCAEODDITMRFPDVISDNVPLPLFHKVP  
TANNSNALVDKFKTSLTGLNFSGVATTE"  
2098..3063  
/locus\_tag="cg2938"  
2098..3063  
/locus\_tag="cg2938"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type dipeptide/oligopeptide/nickel transport  
system, permease component"  
/protein\_id="CAF20677.1"  
/db\_xref="GI:41326833"  
/translation="MSNLRIYVGRRIALPIMITGVTLVPTWSPADPARLAGE  
SASPEALEAYREANGNDPMWRYTFDFILGMLKDGITSSGVAATDIYARFPITLQ  
LTFWGLIAVAVVALIIGVIALALYRDMPOQLRVSIALATPSFMALTLQWGLTI  
PGAMGFEPALVTRWVPESEDPATYFNNNILPAILAVPAGSLARVRSWVEEDKD  
YVRATGAGIPEKTEVVARVNRVALTPTITGLRGVYMGAVIIEIIFNIQAMGOL  
ILDGVNRNDVYLQGVTLVYALAFIIVNLAVDLVLVNPRIKST"  
3064..5067  
/locus\_tag="cg2939"  
3064..5067  
/locus\_tag="cg2939"  
/codon\_start=1  
/transl\_table=11  
/product="ABC-type dipeptide/oligopeptide/nickel transport  
system, fused permease and ATPase components"  
/protein\_id="CAF20678.1"  
/db\_xref="GI:41326834"  
/translation="MRRKTTTLENRGALGFRALAPTSKIALVFLILPLAIFA  
PLAKYDPLASGTPVPPSGEHWGTDAIGDIPSRVAGARASLIIGFATGALLV  
AAVGSIAAANKAISVLMRYLDVIMSPGIALAIPVAVGNGVPTVPAIAPLT  
VFEASLSPFINAGVPPSPWGNLADGKALLLSGAWPFPFGMLTLTVLCNLISE  
GLTDTLASPKPKPVASAKKALKKESGKEGGLVGLHTTREANASLLASIALST  
SENNSNRRLIFDGNPTPLLEBDLKISFPNAGDINIVDGVNPTVAPCGTGLVGESE  
CGKSITMASIMGLIPPAKIEGELIFDGNKLLDKDELALRGHEIMAYODALSLT  
NPSMLISAQKQILTRGGRSABELELVGLDPKRTQLOSYPHEISGGCGRGLVISMAL  
TKNRLIADSEPTALDVTYQOQVVDLNLBKKLGFANITFVSHDLALVRLVHKITV  
MYAGQVQGGTTEIILIDPREHETRLGLSVLSIEGVRLVQVPTVSPKEFVAGD  
RFAPRSEPELIGDQKRVLRPIPTGHAVAATDELAAKGEOR"  
5064..5882  
/locus\_tag="cg2940"  
5064..5882  
/locus\_tag="cg2940"  
/codon\_start=1  
/transl\_table=11  
/product="ATPase components of ABC-type transport system,  
contain duplicated ATPase domains"  
/protein\_id="CAF20679.1"  
/db\_xref="GI:41326835"  
/translation="MTSTIDTRQPIPLVLTANTVIRHTRTKGLFRPDQVHANKINIFKA  
YRDEVTIGVESGCKSTLARVVGLOPTGSGVLEKGMKPRGRORKELESSVSV  
FQDPATSYNRMVTRVREOLDPLRHHVHGVGBADSNQVNLINMVGIPQSGALVPROV  
SGGQRQVATARALAKPDIVADEPTASDVSVRAQVNNILDKTEIGLGLVPIISH  
DINTVRVSDRIAVMLAGEITIEENTTSIFNNAQODYTRTLLEATPSLNTKRL"  
complement (5935..6558)  
/locus\_tag="cg2941"

gene



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:47:34 ; Search time 128.44 Seconds  
(without alignment)  
5630.828 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19

Sequence: 1 tctagcgcgcacacgcgcg 19

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

# SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 19    | 100.0       | 1489   | 9  | AG109756 Pan trogl  |
| 2          | 17    | 89.5        | 163    | 6  | CD323908 StrPu537.  |
| 3          | 17    | 89.5        | 333    | 6  | CD324115 StrPu537.  |
| 4          | 17    | 89.5        | 461    | 6  | CD339324 StrPu536.  |
| 5          | 17    | 89.5        | 570    | 6  | CD912036 G550.113A  |
| 6          | 17    | 89.5        | 676    | 6  | BM607198 170006870  |
| 7          | 17    | 89.5        | 752    | 4  | B1144572 602909436  |
| 8          | 17    | 89.5        | 942    | 7  | CF546859 AGENCOURT  |
| 9          | 16    | 84.2        | 298    | 1  | A1537835 tpe2a08.x  |
| 10         | 16    | 84.2        | 527    | 8  | AQ6464580 HS 5104.A |
| 11         | 16    | 84.2        | 546    | 9  | CE546398 t1gt-gss-  |
| 12         | 16    | 84.2        | 579    | 4  | BF983915 602307473  |
| 13         | 16    | 84.2        | 651    | 9  | AG154721 Pan trogl  |
| 14         | 16    | 84.2        | 1091   | 7  | BF983743 602307337  |
| 15         | 16    | 84.2        | 251    | 7  | CF842669 PSHB021XB  |
| 16         | 15    | 78.9        | 268    | 6  | CD613609 56029053J  |
| 17         | 15    | 78.9        | 278    | 2  | AM428561 66804.MAR  |
| 18         | 15    | 78.9        | 281    | 6  | CD613605 56028933J  |
| 19         | 15    | 78.9        | 410    | 6  | CD613589 56028923J  |
| 20         | 15    | 78.9        | 421    | 2  | BF117366 u236b06.Y  |
| 21         | 15    | 78.9        | 465    | 4  | B1078010 602874382  |
| 22         | 15    | 78.9        | 482    | 5  | BO753281 EBAN01.SQ  |
| 23         | 15    | 78.9        | 484    | 4  | B1339263 364169.MA  |
| 24         | 15    | 78.9        | 508    | 2  | BE227626 894033G09  |

|    |    |      |     |   |            |            |
|----|----|------|-----|---|------------|------------|
| 25 | 15 | 78.9 | 524 | 6 | CB066279   | PVB612B05  |
| 26 | 15 | 78.9 | 544 | 1 | AJ447864   | AJ447864   |
| 27 | 15 | 78.9 | 587 | 2 | BF528035   | BF528035   |
| 28 | 15 | 78.9 | 587 | 6 | CB427482   | CB427482   |
| 29 | 15 | 78.9 | 605 | 4 | BG712523   | BG712523   |
| 30 | 15 | 78.9 | 607 | 5 | BQ038392   | BQ038392   |
| 31 | 15 | 78.9 | 627 | 7 | CF571178   | CF571178   |
| 32 | 15 | 78.9 | 638 | 4 | B1066464   | B1066464   |
| 33 | 15 | 78.9 | 638 | 7 | CN274909   | CN274909   |
| 34 | 15 | 78.9 | 639 | 4 | B1851715   | B1851715   |
| 35 | 15 | 78.9 | 680 | 9 | BX905787   | BX905787   |
| 36 | 15 | 78.9 | 680 | 9 | LBAP021A05 | LBAP021A05 |
| 37 | 15 | 78.9 | 682 | 1 | AJ452357   | AJ452357   |
| 38 | 15 | 78.9 | 695 | 4 | BG764608   | BG764608   |
| 39 | 15 | 78.9 | 715 | 6 | BY707912   | BY707912   |
| 40 | 15 | 78.9 | 723 | 6 | CB457166   | CB457166   |
| 41 | 15 | 78.9 | 730 | 5 | B0251290   | B0251290   |
| 42 | 15 | 78.9 | 736 | 6 | CD342162   | CD342162   |
| 43 | 15 | 78.9 | 758 | 4 | B1156664   | B1156664   |
| 44 | 15 | 78.9 | 762 | 7 | CO879708   | CO879708   |
| 45 | 15 | 78.9 | 771 | 4 | B1660197   | B1660197   |

## ALIGNMENTS

|            |   |             |                 |        |                 |
|------------|---|-------------|-----------------|--------|-----------------|
| RESULT 1   | AG109756  | 1489 bp     | DNA             | linear | GSS 03-NOV-2001 |
| LOCUS      | Pan troglodytes DNA, clone: PTB-115E10.R, genomic survey sequence.  |             |                 |        |                 |
| DEFINITION | AG109756  |             |                 |        |                 |
| ACCESSION  | AG109756.1  | GI:16730275 |                 |        |                 |
| VERSION    | GSS.  |             |                 |        |                 |
| KEYWORDS   | Pan troglodytes (chimpanzee)  |             |                 |        |                 |
| SOURCE     | Pan troglodytes   |             |                 |        |                 |
| ORGANISM   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  |             |                 |        |                 |
| REFERENCE  | 1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  |             |                 |        |                 |
| AUTHORS    | Unpublished   |             |                 |        |                 |
| TITLE      | BAC end sequences of library PTB  |             |                 |        |                 |
| JOURNAL    | 2 (bases 1 to 1489)   |             |                 |        |                 |
| REFERENCE  | Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@cc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170) |             |                 |        |                 |
| AUTHORS    | Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  |             |                 |        |                 |
| TITLE      | PRIMERS   |             |                 |        |                 |
| JOURNAL    | Sequencing: M13rev  |             |                 |        |                 |
| REFERENCE  | Vector : pKS145   |             |                 |        |                 |
| AUTHORS    | R.Site 1 : SacI   |             |                 |        |                 |
| TITLE      | R.Site 2 : SacI   |             |                 |        |                 |
| JOURNAL    | Location/Qualifiers   |             |                 |        |                 |
| FEATURES   | 1..1489   |             |                 |        |                 |
| SOURCE     | /organism="Pan troglodytes"   |             |                 |        |                 |
|            | /mol_type="genomic DNA"   |             |                 |        |                 |
|            | /db_xref="taxon:9598"   |             |                 |        |                 |
|            | /clone="PTB-115E10.R"   |             |                 |        |                 |
|            | /sex="male"   |             |                 |        |                 |
|            | /cell_type="lymphoblast"  |             |                 |        |                 |
|            | /clone_id="PTB Chimpanzee Male BAC library"   |             |                 |        |                 |
| ORIGIN     | Query Match   | 100.0%;     | Score 19;       | DB 9;  | Length 1489;    |
|            | Best Local Similarity   | 100.0%;     | Pred. No. 0.27; |        |                 |

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGT 19  
|||||

Db 739 TGTAGCGGCCACGCGT 757

RESULT 2 CD323908 163 bp mRNA linear EST 17-SEP-2003  
LOCUS STpU537.002270 Sea urchin embryo 20hr blastula stage cDNA library  
DEFINITION MPMG537 Strongylocentrotus purpuratus cDNA clone  
VERSION CALTP537K1047;MPI\_537\_17E12 3', mRNA sequence.  
KEYWORDS CD323908.1 GI:34795969  
SOURCE EST.  
ORGANISM Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidea; Euechinoidea; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
REFERENCE 1 (bases 1 to 163)  
Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,  
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.  
Generation, annotation, evolutionary analysis, and database  
integration of 20,000 unique sea urchin EST clusters  
Genome Res. 13 (12), 2736-2746 (2003)  
JOURNAL Contact: Poustka AJ  
COMMENT laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr. 63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: poustka@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONF) to reduce sequencing redundancy. According to the ONF  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per ONF cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well as the coordinates of  
the other clones assigned to the same ONF cluster as the clone from  
which the above EST is generated is available at the sea urchin  
project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA  
clones and filters are distributed via the Resource Center/Primary  
Database of the German Human Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGGCTTACACTTATGCTTCCGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGCGACCTGCGGAAGGGGATGTG 3' (M13FSP) 3'-seq  
Seq primer: 5' GCTATTACGCGACCTGCGGAAGGGGATGTG 3' (M13FSP)  
High quality sequence stop: 163.  
Location/Qualifiers  
1. 163  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="CALTP537K1047;MPI\_537\_17E12"  
/issue\_type="whole embryo"  
/tissue\_type="embryonic 20hr"  
/dev\_stage="embryonic 20hr"  
/lab\_host="E.coli, XL1 blue"  
/clone\_1lb="Sea urchin embryo 20hr blastula stage cDNA  
library MPMG537"  
/note="Vector: pSport1; site 1: NotI; site 2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-TGACTAGTTCTAGATCGGACGCGCCG (T)15-3' and a  
SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL) "

ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 163;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGT 17  
|||||

Db 135 TGTAGCGGCCACGCGT 151  
|||||

RESULT 3 CD324115/c 333 bp mRNA linear EST 17-SEP-2003  
LOCUS STpU537.002477 Sea urchin embryo 20hr blastula stage cDNA library  
DEFINITION MPMG537 Strongylocentrotus purpuratus cDNA clone  
VERSION CALTP537K1047;MPI\_537\_47K10 3', mRNA sequence.  
KEYWORDS CD324115  
CD324115 GI:34796176  
SOURCE EST.  
ORGANISM Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidea; Euechinoidea; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
REFERENCE 1 (bases 1 to 333)  
Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,  
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.  
Generation, annotation, evolutionary analysis, and database  
integration of 20,000 unique sea urchin EST clusters  
Genome Res. 13 (12), 2736-2746 (2003)  
JOURNAL Contact: Poustka AJ  
COMMENT laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr. 63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: poustka@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONF) to reduce sequencing redundancy. According to the ONF  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per ONF cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well as the coordinates of  
the other clones assigned to the same ONF cluster as the clone from  
which the above EST is generated is available at the sea urchin  
project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA  
clones and filters are distributed via the Resource Center/Primary  
Database of the German Human Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGGCTTACACTTATGCTTCCGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGCGACCTGCGGAAGGGGATGTG 3' (M13FSP) 3'-seq  
Seq primer: 5' GCTATTACGCGACCTGCGGAAGGGGATGTG 3' (M13FSP)  
High quality sequence stop: 333.  
Location/Qualifiers  
1. 333  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="CALTP537K1047;MPI\_537\_47K10"  
/issue\_type="whole embryo"  
/tissue\_type="embryonic 20hr"  
/dev\_stage="embryonic 20hr"  
/lab\_host="E.coli, XL1 blue"  
/clone\_1lb="Sea urchin embryo 20hr blastula stage cDNA  
library MPMG537"  
/note="Vector: pSport1; site 1: NotI; site 2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-TGACTAGTTCTAGATCGGACGCGCCG (T)15-3' and a  
SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL) "

ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 333;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGT 17  
|||||

Db 30 TGTAGCGGCCACGCGT 14  
|||||

RESULT 4  
LOCUS  
CD339324 461 bp mRNA linear EST 17-SEP-2003  
DEFINITION  
StrP536.003145 Sea urchin embryo 40hr gastrula stage cDNA library  
MPM6536 Strongylocentrotus purpuratus cDNA clone  
CALP536H0915.MPI\_536\_15H9 3', mRNA sequence.  
ACCESSION  
CD339324  
VERSION  
CD339324.1 GI:34805850  
SOURCE  
Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidea; Euechinoidea; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
REFERENCE  
1 (bases 1 to 461)  
Pousatka, A., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,  
Reinhardt, R., Herwig, R., Panopoulou, G. and Lehrach, H.  
Generation, annotation, evolutionary analysis, and database  
integration of 20,000 unique sea urchin EST clusters  
Genome Res. 13 (12), 2736-2746 (2003)  
JOURNAL  
Contact: Pousatka, A.  
laboraty 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Imestr. 63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: pousatka@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONP) to reduce sequencing redundancy. According to the ONP  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per ONP cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well as the coordinates of  
which the above EST is generated is available at the sea urchin  
project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA  
clones and filters are distributed via the Resource Center/Primary  
Database of the German Human Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGGCTTACCTTATGCTTCGGGCTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGCCAGCTGCGAAGGGGAGTGTG 3' (M13FSP) 3'-seq  
Seq primer: 5' GCTATTACGCCAGCTGCGAAGGGGAGTGTG 3' (M13FSP)  
High quality sequence stop: 461.  
location/Qualifiers  
1..461  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="CALTP536H0915.MPI\_536\_15H9"  
/issue="whole embryo"  
/dev\_stage="embryonic 40hr"  
/lab\_host="E. coli, Xtl blue"  
/clone\_lib="Sea urchin embryo 40hr gastrula stage cDNA  
library MPM6536"  
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Random  
primed and directionally cloned in pSport1 vector using a  
NotI (5'-GGACTAGTCTTAGATCGGAGGAGGGCGCC (T)15-3' and a  
SalI 5'-TGACCCACGCGTCCG-3' adapters (Gibco BRL)."

## ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 461;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGGCCACGCGT 17  
|||||  
Db 290 TGTAGCGGCCACGCGT 306

RESULT 5  
CD912036/c

LOCUS  
CD912036 570 bp mRNA linear EST 14-JUL-2003  
DEFINITION  
G550.113A12P010522 G550 Triticum aestivum cDNA clone G550113A12,  
mRNA sequence.  
ACCESSION  
CD912036  
VERSION  
CD912036.1 GI:32686360  
KEYWORDS  
EST.  
SOURCE  
Triticum aestivum (bread wheat)  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
REFERENCE  
1 (bases 1 to 570)  
Genoplate.  
Unpublished (2003)  
JOURNAL  
Contact: Genoplate  
Genoplate  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplate' (<http://www.genoplate.com>  
and <http://genoplate-info.inbioogen.fr>).  
location/Qualifiers  
1..570  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="rectal"  
/db\_xref="taxon:4565"  
/clone="G550113A12"  
/issue\_type="grain (550 degrees per day after  
pollination)"  
/clone\_lib="G550"

## ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 570;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGGCCACGCGT 17  
|||||  
Db 20 TGTAGCGGCCACGCGT 4

## FEATURES

## source

RESULT 6  
BM607198/c 676 bp mRNA linear EST 25-FEB-2002  
LOCUS  
DEFINITION  
17000687084830 A.Gam.ad.cDNA.blod1 Anopheles gambiae cDNA clone  
19600449683177 5', mRNA sequence.  
ACCESSION  
BM607198  
VERSION  
BM607198.1 GI:18905302  
KEYWORDS  
EST.  
SOURCE  
Anopheles gambiae (African malaria mosquito)  
ORGANISM  
Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicidae;  
Anopheles.  
1 (bases 1 to 676)  
Holt, R.A., Jin, J.-Y., Murphy, S.D., Evans, C.A., Kraft, C.L.,  
Charlab, R., Collins, F.H., Venter, J.C. and Hoffman, S.L.  
Celera Anopheles gambiae EST project  
Unpublished (2002)  
JOURNAL  
Contact: Holt R.A.  
Celera Genomics  
45 W. Gude Dr., Rockville, MD 20850, USA  
Tel: 2404533151  
Fax: 2404534580  
Email: [holt@celera.com](mailto:holt@celera.com)  
Plate: N001004AS row: L column: 11  
Seq primer: M13 Reverse.  
location/Qualifiers  
1..676  
/organism="Anopheles gambiae"

## ORIGIN

Query Match 89.5%; Score 17; DB 6; Length 570;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTAGCGGCCACGCGT 17  
|||||  
Db 20 TGTAGCGGCCACGCGT 4

RESULT 5  
CD912036/c

/mol\_type="mRNA"  
/strain="RSP-ST (Reduced susc. to Permethrin - std.  
chromosome)"  
/db\_xref="taxon:7165"  
/clone="19600449683177"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Whole  
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24  
hours after human blood feeding. cDNA inserts >500 bp  
cloned directionally into pSPORT 1. Not 1 site 18 3'.  
Clones available through the Malaria Research and  
Reference Reagent Resource Center (www.malaria.mr4.org)"

ORIGIN  
Query Match 89.5%; Score 17; DB 4; Length 676;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGACGCGT 17  
|||||  
Db 56 TGTAGCGGCCCGACGCGT 40

RESULT 7 752 bp mRNA linear EST 05-JUL-2001  
B1144572/c 602909436F1 NCI\_CGAP\_L19 Mus musculus cDNA clone IMAGE:5050424 5',  
LOCUS mRNA sequence.  
DEFINITION B1144572  
ACCESSION B1144572  
VERSION B1144572.1 GI:14604573  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 752)  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Found through the I.M.A.G.E. Consortium information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LLAM1137 row: h column: 09  
High quality sequence stop: 445.  
Location/Qualifiers  
1..752  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5050424"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.9 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

FEATURES  
source

ORIGIN

Query Match 89.5%; Score 17; DB 4; Length 752;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGACGCGT 17  
|||||

Db 19 TGTAGCGGCCCGACGCGT 3  
RESULT 8 942 bp mRNA linear EST 22-SEP-2003  
CF546859/c A6ENECOURT\_15607895 NICHG\_XGC\_Kid1 Xenopus laevis cDNA clone  
LOCUS IMAGE:7008145 5', mRNA sequence.  
DEFINITION CF546859  
ACCESSION CF546859.1 GI:34883691  
VERSION EST.  
KEYWORDS Xenopus laevis (African clawed frog)  
SOURCE Xenopus laevis  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Xenopus.  
1 (bases 1 to 942)  
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. Igor David  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LLAM14711 row: g column: 23  
High quality sequence stop: 642.  
Location/Qualifiers  
1..942  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:7008145"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.2 kb. Constructed by Life  
Technologies. Note: This is a Xenopus Gene Collection  
(XGC) library."

FEATURES  
source

ORIGIN

Query Match 89.5%; Score 17; DB 7; Length 942;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGACGCGT 17  
|||||  
Db 32 TGTAGCGGCCCGACGCGT 16

RESULT 9 298 bp mRNA linear EST 13-MAY-1999  
A1537835/c tp26a08.x1 NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE:218886 3',  
LOCUS similar to TR:Q13488 Q13488 SPECIFIC 116-KDA VACUOLAR PROTON PUMP  
DEFINITION SUBMIT. :, mRNA sequence.  
ACCESSION A1537835  
VERSION A1537835.1 GI:4451970  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 298)



**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
**JOURNAL** Tumor Gene Index  
**COMMENT** Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-rc@mail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
[www.bio.lnl.gov/bdtp/image/image.html](http://www.bio.lnl.gov/bdtp/image/image.html)  
 Insert length: 1961 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 168  
 POLY-A-NO.

# FEATURES

source  
 1. .298  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2188886"  
 /tissue\_type="poorly differentiated adenocarcinoma with  
 signed ring cell features"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Gas4"  
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

# ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 GTAGCGGCCCGCGCT 17  
 |||||  
 Db 244 GTAGCGGCCCGCGCT 229

**RESULT 10** A0464580 527 bp DNA linear GSS 23-APR-1999  
**LOCUS** HS\_5104\_A1\_G12\_T7A\_RPCT-11 Human Male BAC Library Homo sapiens  
**DEFINITION** genomic clone Plate=680 Col=23 Row=M, genomic survey sequence.  
**ACCESSION** A0464580  
**VERSION** A0464580.1 GI:4641675  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
**AUTHORS** Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 527)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
 Hood, L.  
**TITLE** Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
**MEDLINE** 99380589  
**PUBMED** 10449764  
**COMMENT** Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCT-11. For BAC  
 library availability, please contact Pieter de Jong

(pieterdejong.med.bufiled.edu). Clones may be purchased from  
 BACPAC Resources ([http://bacpac.med.bufiled.edu/ordering\\_bac.htm](http://bacpac.med.bufiled.edu/ordering_bac.htm))  
 or from Research Genetics (info@resgen.com). BAC end Web Server:  
<http://www.htsc.washington.edu>  
 Plate: 680 row: M column: 23  
 Seq primer: 17  
 Class: BAC ends  
 High quality sequence stop: 527.  
 Location/Qualifiers

source  
 1. .527  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=680 Col=23 Row=M"  
 /sex="male"  
 /clone\_lib="RPCT-11 Human Male BAC Library"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI methylase. Size selected DNA was cloned into the  
 pBAC3.6 vector at EcoRI sites"

# ORIGIN

Query Match 84.2%; Score 16; DB 8; Length 527;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 4 AGCGGCCCGCGCTGG 19  
 |||||  
 Db 24 AGCGGCCCGCGCTGG 9

**RESULT 11** CE548398 546 bp DNA linear GSS 28-SEP-2003  
**LOCUS** tigr-gss-dog-17000327426188 Dog Library Canis familiaris genomic,  
**DEFINITION** genomic survey sequence.  
**ACCESSION** CE548398  
**VERSION** CE548398.1 GI:36865179  
**KEYWORDS** GSS.  
**SOURCE** Canis familiaris (dog)  
**ORGANISM** Canis familiaris  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
**AUTHORS** Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 1 (bases 1 to 546)  
 Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,  
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
 Venter, J.C.  
**TITLE** The dog genome: survey sequencing and comparative analysis  
**JOURNAL** Science 301 (5641), 1898-1903 (2003)  
**MEDLINE** 14512627  
**PUBMED** 14512627  
**COMMENT** Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirkness@tigr.org  
 Class: shotgun.  
**FEATURES** Location/Qualifiers  
 source  
 1. .546  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_lib="Dog Library"  
 /note="Site\_1: BstXI; Libraries were prepared from  
 peripheral blood"

# ORIGIN

Query Match 84.2%; Score 16; DB 9; Length 546;  
 Best Local Similarity 100.0%; Pred. No. 20;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCGACGCGTGG 19  
|||||  
268 AGCGGCCCGACGCGTGG 253

Db

RESULT 12  
BP983915/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT

BP983915 579 bp mRNA linear EST 23-JAN-2001  
602307473F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4398619 5',  
mRNA sequence.  
BP983915  
BP983915.1 GI:12386727  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LRAM10101 row: a column: 20  
High quality sequence stop: 579.  
Location/Qualifiers

FEATURES  
source  
1..579  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4398619"  
/tissue\_type="nododenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site 1: NCBI; Site 2: Sali; Cloned unidirectionally;  
oligo-dt primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 84.2%; Score 16; DB 4; Length 579;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCCGACGCGT 17  
|||||  
17 GTAGCGGCCCGACGCGT 2

Db

RESULT 13  
AG154721/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AG154721 651 bp DNA linear GSS 09-JAN-2002  
Pan troglodytes DNA, clone: RP43-018N05.T7, genomic survey  
sequence.  
AG154721  
AG154721.1 GI:16684399  
GSS.  
Pan troglodytes (chimpanzee)  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT

Totoki, Y., Watanabe, H. and Sakaki, Y.  
BAC end sequences of library RPCT-43  
Unpublished  
2 (bases 1 to 651)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suenho-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library RPCT-43 This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
PRIMERS  
Sequencing: T7  
LIBRARY  
Vector : pBAC3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.  
Location/Qualifiers

FEATURES  
source  
1..651  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-018N05.T7"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RPCT-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 84.2%; Score 16; DB 9; Length 651;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGCGGCCCGACGCGTGG 19  
|||||  
500 AGCGGCCCGACGCGTGG 485

Db

RESULT 14  
BP983743/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT

BP983743 1091 bp mRNA linear EST 23-JAN-2001  
602307337F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4398600 5',  
mRNA sequence.  
BP983743  
BP983743.1 GI:12386555  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LRAM10101 row: 1 column: 09  
High quality sequence stop: 689.  
Location/Qualifiers

FEATURES  
source  
1..1091  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

Search completed: February 10, 2005, 17:01:53  
 Job time : 132.44 secs

## ORIGIN

/clone="IMAGE:4398800"  
 /issue\_type="duodenal adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 88"  
 /note="Organ: small\_intestine; Vector: PCMV-SPORT6;  
 Site\_1: Not; Site\_2: SalI; Cloned unidirectionally;  
 oligo-dt primed. Average insert size 1.767 kb. Library  
 enriched for full-length clones and constructed by Life  
 Technologies. Note: this is a NIH\_MGC Library."

Query Match 84.2%; Score 16; DB 4; Length 1091;  
 Best Local Similarity 100.0%; Pred.No.19;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGCCACCGCT 17  
 |||||  
 Db 17 GTAGCGCCACCGCT 2

## RESULT 15

CF842669

LOCUS

DEFINITION CF842669 251 bp mRNA linear EST 30-OCT-2003  
 PSIB021XB206 USDA-IFARS:Expression of Phytophthora sojae genes  
 during infection and propagation\_SHB Phytophthora sojae cDNA clone

SHB021B20 5, mRNA sequence.  
 CF842669

CF842669.1 GI:38058323  
 EST.

PHYTOPHTHORA sojae  
 PHYTOPHTHORA sojae

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;  
 Phytophthora.

1 (bases 1 to 251)

REFERENCE Tyler,B. Not Published

TITLE Tyler,B. Not Published

JOURNAL Unpublished (2003)

COMMENT Contact: Tyler B

TYLER lab

VBI

1880 Pratt Dr., Blackebury, VA 24061, USA

Tel: 540-231-7318

Email: dmtyler@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 021 row: B column: 20

Seq primer: BK reverse primer

High quality sequence stop: 251.

location/Qualifiers

1..251

/organism="Phytophthora sojae"

/mol\_type="mRNA"

/db\_xref="taxon:67593"

/clone="SHB021B20"

/issue\_type="mycelium"

/cell\_line="P6497"

/dev\_stage="48 hr. post infection stage"

/lab\_host="Soybean plant"

/clone\_lib="USDA-IFARS:Expression of Phytophthora sojae  
 genes during infection and propagation\_SHB"

/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match

Best Local

Matches 15; Conservative

QY 5 GCGGCCACGCGTGG 19

Db 62 GCGGCCACGCGTGG 76

78.9%; Score 15; DB 7; Length 251;

Similarity 100.0%; Pred.No. 91; Indels 0; Gaps 0;

Mismatches 0;

GCGGCCACGCGTGG 76

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2005, 07:18:58 ; Search time 5.8365 Seconds  
(without alignments)  
5324.730 Million cell updates/sec

Title: US-10-790-430-9

Perfect score: 19  
Sequence: 1 tctagcgccaccacgctgg 19

Scoring table: OLIGO\_NTC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCUS.COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                 | Description        |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1          | 19    | 100.0       | 19     | US-09-872-051-9    | Sequence 9, Appl1  |
| 2          | 19    | 100.0       | 498    | US-09-872-051-7    | Sequence 7, Appl1  |
| 3          | 15    | 78.9        | 843    | US-09-902-540-9093 | Sequence 9093, Ap  |
| 4          | 15    | 78.9        | 1405   | US-08-390-162-3    | Sequence 3, Appl1  |
| 5          | 15    | 78.9        | 1405   | US-08-685-945B-3   | Sequence 3, Appl1  |
| 6          | 15    | 78.9        | 2107   | US-08-390-162-1    | Sequence 1, Appl1  |
| 7          | 15    | 78.9        | 2107   | US-08-685-945B-1   | Sequence 1, Appl1  |
| 8          | 15    | 78.9        | 2558   | US-09-902-540-6342 | Sequence 6342, Ap  |
| 9          | 15    | 78.9        | 2560   | US-09-902-540-400  | Sequence 400, App  |
| 10         | 15    | 78.9        | 10915  | US-09-902-540-972  | Sequence 972, App  |
| 11         | 14    | 73.7        | 1083   | US-09-902-540-5440 | Sequence 5440, Ap  |
| 12         | 14    | 73.7        | 1092   | US-09-902-540-8179 | Sequence 8179, Ap  |
| 13         | 14    | 73.7        | 1209   | US-09-826-509-556  | Sequence 556, App  |
| 14         | 14    | 73.7        | 1376   | US-09-016-434-1361 | Sequence 1361, Ap  |
| 15         | 14    | 73.7        | 1394   | US-08-068-729-3    | Sequence 3, Appl1  |
| 16         | 14    | 73.7        | 1394   | US-09-255-671-3    | Sequence 3, Appl1  |
| 17         | 14    | 73.7        | 1394   | US-09-395-366-3    | Sequence 3, Appl1  |
| 18         | 14    | 73.7        | 1488   | US-09-096-776B-3   | Sequence 3, Appl1  |
| 19         | 14    | 73.7        | 1488   | US-09-923-922-3    | Sequence 3, Appl1  |
| 20         | 14    | 73.7        | 2457   | US-09-949-016-1330 | Sequence 1330, Ap  |
| 21         | 14    | 73.7        | 2457   | US-09-949-016-1331 | Sequence 1331, Ap  |
| 22         | 14    | 73.7        | 2534   | US-09-096-776B-1   | Sequence 1, Appl1  |
| 23         | 14    | 73.7        | 2534   | US-09-923-922-1    | Sequence 1, Appl1  |
| 24         | 14    | 73.7        | 2640   | US-08-684-932A-37  | Sequence 37, Appl1 |
| 25         | 14    | 73.7        | 2640   | US-09-618-304B-1   | Sequence 2404, Ap  |
| 26         | 14    | 73.7        | 2654   | US-09-949-016-2404 | Sequence 2405, Ap  |
| 27         | 14    | 73.7        | 2654   | US-09-949-016-2405 | Sequence 2405, Ap  |

|      |    |      |       |   |                     |                    |
|------|----|------|-------|---|---------------------|--------------------|
| C 28 | 14 | 73.7 | 2655  | 4 | US-09-016-434-1094  | Sequence 1094, Ap  |
| C 29 | 14 | 73.7 | 2655  | 4 | US-09-023-655-916   | Sequence 916, App  |
| C 30 | 14 | 73.7 | 2706  | 4 | US-09-949-016-2975  | Sequence 2975, Ap  |
| C 31 | 14 | 73.7 | 2706  | 4 | US-09-949-016-2976  | Sequence 2976, Ap  |
| C 32 | 14 | 73.7 | 3465  | 4 | US-09-489-039A-553  | Sequence 553, App  |
| C 33 | 14 | 73.7 | 3603  | 4 | US-09-902-540-3466  | Sequence 3266, App |
| C 34 | 14 | 73.7 | 11917 | 4 | US-09-949-016-13072 | Sequence 13072, A  |
| C 35 | 14 | 73.7 | 11917 | 4 | US-09-949-016-13073 | Sequence 13073, A  |
| C 36 | 14 | 73.7 | 13256 | 4 | US-09-902-540-1006  | Sequence 1006, Ap  |
| C 37 | 14 | 73.7 | 14638 | 4 | US-09-902-540-1106  | Sequence 1106, Ap  |
| C 38 | 14 | 73.7 | 15061 | 4 | US-09-949-016-14717 | Sequence 14717, A  |
| C 39 | 14 | 73.7 | 15061 | 4 | US-09-949-016-14718 | Sequence 14718, A  |
| C 40 | 14 | 73.7 | 15859 | 4 | US-09-949-016-14146 | Sequence 14146, A  |
| C 41 | 14 | 73.7 | 15859 | 4 | US-09-949-016-14147 | Sequence 14147, A  |
| C 42 | 14 | 73.7 | 19954 | 4 | US-09-902-540-1150  | Sequence 1150, Ap  |
| C 43 | 14 | 73.7 | 36800 | 4 | US-08-311-731A-139  | Sequence 139, App  |
| C 44 | 14 | 73.7 | 41310 | 4 | US-09-902-540-1264  | Sequence 1264, Ap  |
| C 45 | 13 | 68.4 | 36    | 3 | US-08-434-099A-36   | Sequence 36, Appl1 |

#### ALIGNMENTS

RESULT 1  
US-09-872-051-9  
; Sequence 9, Application US/09872051  
; Patent No. 6825400  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine  
; APPLICANT: Heck, Gregory  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: Corn Event PV-ZMG732(nk603) and Composition and Methods for Det  
; FILE REFERENCE: 38-21(52258)B  
; CURRENT APPLICATION NUMBER: US/09/872,051  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/213,567  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/241,215  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/240,014  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: source  
; LOCATION: (1)..(19)  
; OTHER INFORMATION: Zea maize genomic and vector DNA  
US-09-872-051-9  
Query Match 100.0%; Score 19; DB 4; Length 19;  
Best local Similarity 100.0%; Pred. No. 0.16;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TGTAGCGCCACGCGTGG 19  
DB 1 TGTAGCGCCACGCGTGG 19  
RESULT 2  
US-09-872-051-7  
; Sequence 7, Application US/09872051  
; Patent No. 6825400  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Co  
; APPLICANT: Behr, Carl  
; APPLICANT: Hironaka, Catherine

APPLICANT: Heck, Gregory  
APPLICANT: You, Jinsong  
TITLE OF INVENTION: Corn Event PV-ZMG732(hk603) and Composition and Methods for Detecting  
TITLE OF INVENTION: Thereof  
FILE REFERENCE: 38-21(52258)B  
CURRENT APPLICATION NUMBER: US/09/872,051  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: 60/213,567  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/241,215  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/240,014  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 7  
LENGTH: 498  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)-(498)  
OTHER INFORMATION: 1-304 Zea maize genomic DNA  
OTHER INFORMATION: 305-349 construct vector DNA  
OTHER INFORMATION: 350-498 rice actin 1 promoter DNA  
US-09-872-051-7

Query Match 100.0%; Score 19; DB 4; Length 498;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGTGG 19  
DB 296 TGTAGCGGCCACGCGTGG 314

RESULT 3  
US-09-902-540-9093/C  
Sequence 9093, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(115849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 9093  
LENGTH: 843  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
US-09-902-540-9093

Query Match 78.9%; Score 15; DB 4; Length 843;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTAGCGGCCACGCGC 16  
DB 693 GTAGCGGCCACGCGC 679

RESULT 4  
US-08-390-162-3/C  
Sequence 3, Application US/08390162  
Patent No. 5576192  
GENERAL INFORMATION:  
APPLICANT: Ichikawa, Atsushi

APPLICANT: Natumiya, Shuh  
TITLE OF INVENTION: Proctoglandin E Receptors, Their DNA and  
TITLE OF INVENTION: Production  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,162  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/024179  
FILING DATE: 23-FEB-1993  
APPLICATION NUMBER: JP 036580-1992  
FILING DATE: 24-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 064889-1992  
FILING DATE: 23-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordie, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04221-0020-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1405 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-390-162-3

Query Match 78.9%; Score 15; DB 1; Length 1405;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCACGCGC 15  
DB 631 TGTAGCGGCCACGCGC 617

RESULT 5  
US-08-685-945B-3/C  
Sequence 3, Application US/08685945B  
Patent No. 5804415  
GENERAL INFORMATION:  
APPLICANT: Ichikawa, Atsushi  
APPLICANT: Natumiya, Shuh  
TITLE OF INVENTION: Proctoglandin E Receptors, Their DNA and  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,945B  
FILING DATE: 22-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/024179  
FILING DATE: 23-FEB-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 036580-1992  
FILING DATE: 24-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 064889-1992  
FILING DATE: 23-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04221-0020-02000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1405 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-685-945B-3

Query Match 78.9%; Score 15; DB 1; Length 1405;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGC 15  
Db 631 TGTAGCGGCCCGC 617

RESULT 6  
US-08-390-162-1/c  
Sequence 1, Application US/08390162  
Patent No. 5576192  
GENERAL INFORMATION:  
APPLICANT: Ichikawa, Atsushi  
APPLICANT: Natumiya, Shuh  
TITLE OF INVENTION: Prostaglandin E Receptors, Their DNA and  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,162  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/024179  
FILING DATE: 23-FEB-1993  
APPLICATION NUMBER: JP 036580-1992  
FILING DATE: 24-FEB-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 064889-1992  
FILING DATE: 23-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04221-0020-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2107 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-390-162-1

Query Match 78.9%; Score 15; DB 1; Length 2107;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTAGCGGCCCGC 15  
Db 644 TGTAGCGGCCCGC 630

RESULT 7  
US-08-685-945B-1/c  
Sequence 1, Application US/08685945B  
Patent No. 5804415  
GENERAL INFORMATION:  
APPLICANT: Ichikawa, Atsushi  
APPLICANT: Natumiya, Shuh  
TITLE OF INVENTION: Prostaglandin E Receptors, Their DNA and  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,945B  
FILING DATE: 22-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/024179  
FILING DATE: 23-FEB-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 036580-1992  
FILING DATE: 24-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 064889-1992  
FILING DATE: 23-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04221-0020-02000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2107 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-685-945B-1

Query Match 78.9%; Score 15; DB 1; Length 2107;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTAGCGGCCCGCGC 15  
Db 644 TGTAGCGGCCCGCGC 630

## RESULT 8

US-09-902-540-6342/C  
Sequence 6342, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 6342  
LENGTH: 2558  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(2558)  
OTHER INFORMATION: unsure at all n locations  
US-09-902-540-6342

Query Match 78.9%; Score 15; DB 4; Length 2558;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTAGCGGCCCGCGC 16  
Db 162 GTAGCGGCCCGCGC 148

## RESULT 9

US-09-902-540-400/C  
Sequence 400, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 400  
LENGTH: 2560  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(2560)

OTHER INFORMATION: unsure at all n locations  
US-09-902-540-400

Query Match 78.9%; Score 15; DB 4; Length 2560;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTAGCGGCCCGCGC 16  
Db 164 GTAGCGGCCCGCGC 150

## RESULT 10

US-09-902-540-972  
Sequence 972, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 972  
LENGTH: 10915  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
US-09-902-540-972

Query Match 78.9%; Score 15; DB 4; Length 10915;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTAGCGGCCCGCGC 16  
Db 152 GTAGCGGCCCGCGC 166

## RESULT 11

US-09-902-540-5440/C  
Sequence 5440, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 5440  
LENGTH: 1083  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
US-09-902-540-5440

Query Match 73.7%; Score 14; DB 4; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TAGCGGCCCGCGC 16  
Db 355 TAGCGGCCCGCGC 342



```
RESULT 12
US-09-902-540-8179/c
; Sequence 8179, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-17,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8179
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8179

Query Match      73.7%; Score 14; DB 4; Length 1092;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 GTAGCGGCCACGCG 15
Db      921 GTAGCGGCCACGCG 908

RESULT 13
US-09-826-509-556/c
; Sequence 556, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Lin, I-Lin
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; FILE REFERENCE: AREN-207
; CURRENT FILING DATE: 2001-04-05
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/1170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 556
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-826-509-556

Query Match      73.7%; Score 14; DB 4; Length 1209;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 TAGCGGCCACGCG 16
Db      536 TAGCGGCCACGCG 523

RESULT 14
US-09-016-434-1361/c
; Sequence 1361, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seihamer
```

```
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; FILE REFERENCE: 1490
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT FILING DATE: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1361:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g410208
US-09-016-434-1361

Query Match      73.7%; Score 14; DB 4; Length 1376;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 TAGCGGCCACGCG 16
Db      610 TAGCGGCCACGCG 597

RESULT 15
US-08-068-729-3/c
; Sequence 3, Application US/08068729
; Patent No. 5985597
; GENERAL INFORMATION:
; APPLICANT: Ford-Hutchinson, Anthony
; APPLICANT: Funk, Colin
; APPLICANT: Grygorczyk, Richard
; APPLICANT: Metters, Kathleen
; TITLE OF INVENTION: DNA Encoding Prostaglandin Receptor EPI
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: JOHN W. WALLIN III
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,729
; FILING DATE: 26-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WALLEN, JOHN W III
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-068-729-3

```

```

Query Match      73.7%; Score 14; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 TAGCGGCCCAAGCG 16
        |||||
Db      610 TAGCGGCCCAAGCG 597

```

Search completed: February 10, 2005, 08:49:48  
 Job time : 6.83865 secs